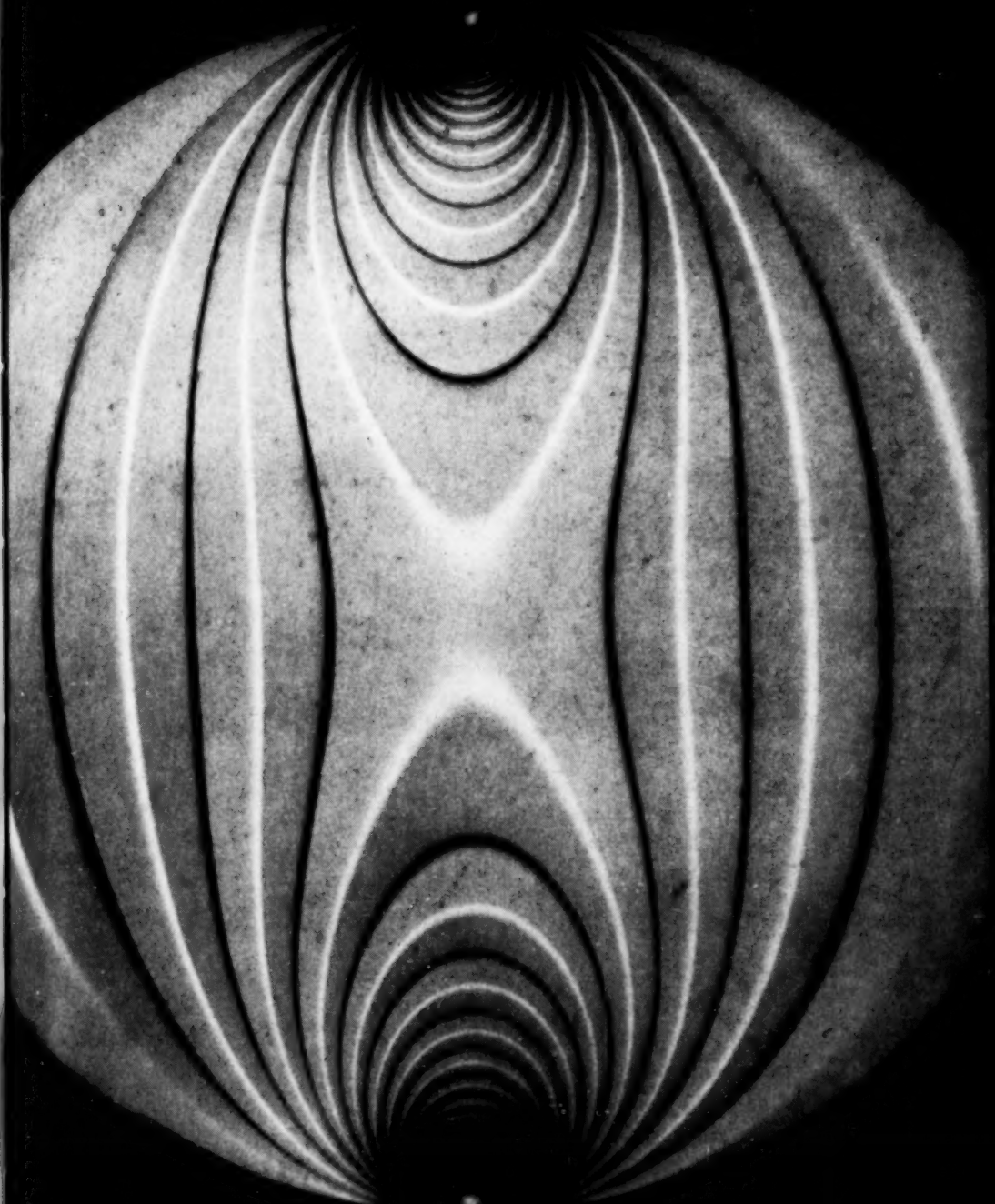
JOURNAL OF APPLIED PHYSICS



JOURNAL OF APPLIED PHYSICS

Vol. 25, No. 8 AUGUST, 1954

In This Issue

		Page
Theorem Concerning Noise in Electron Streams	J. R. Pierce	931
Transfer Function as a Tool in the Analysis of a Resistance-Capacitance Coupled Voltage Amplifier	D. L. Ming and R. W. Thorpe	934
Geometrical Acoustics. I. The Theory of Weak Shock Waves	Joseph B. Keller	938
Transient Wave Analysis in a Linear, Time-Dependent Material	R. D. Glauz and E. H. Lee	947
Thermionic Emission and Electron Diffraction from Thin Films of Barium Oxide	Paul N. Russell and A. S. Eisenstein	954
On the Nature of Radiation Damage in Metals	John A. Brinkman	961
Capillary Viscometer with Continuously Varying Pressure Head	Samuel H. Maron, Irvin M. Krieger, and Arthur W. Sisko	971
Physical Theory of New Circuit Representation for Junction Transistors	Jakob Zawels	976
Experimental Procedures for Determining the Efficiency of Four-Terminal Networks	H. F. Mathis	982
On the Chromatic Field Aberration of the Magnetic Electron Lens in the Electron Microscope	Nozomu Morito	986
Statistical Hydrodynamics in Porous Media	Adrian E. Scheidegger	994
Coincidence of Pulse Trains	Henry D. Friedman	1001
Joint Distribution of n Successive Outputs of a Linear Detector	William C. Hoffman	1006
Application of Ideal Gas Theory to the Gaseous Expansion from an Electric Spark	R. B. Edmonson, H. L. Olsen, and E. L. Gayhart	1008
Factors Determining the Permanent Magnet Properties of Single Crystals of Fe_2NiAl	E. A. Nesbitt, H. J. Williams, and R. M. Bozorth	1014
Electrodynamic Perturbation Theorem, with Application to Nonreciprocal Systems	Alfred G. Redfield	1021
Short-Time Frequency Measurement of Narrow-Band Random Signals in the Presence of Wide-Band Noise	Peter M. Schultheiss, Conrad A. Wogrin, and Felix Zweig	1025
On the Distributions of Signals and Noise after Rectification and Filtering	M. A. Meyer and David Middleton	1037
Letters to the Editor:		
Diffraction of Electromagnetic Waves by an Aperture in a Large Screen	L. Lewin	1053
Electron Atmosphere in Dielectrics	M. J. Morant	1053
Electrical Phenomena in Adhesion. I. Electron Atmospheres in Dielectrics	D. C. West	1054
Electrical Phenomena in Adhesion: Further Comments on Electron Atmospheres	Selby M. Skinner, Robert L. Savage, and John E. Rutzler, Jr.	1055
Direct-Current Transients in Polymethyl Methacrylate and in Polystyrene	Paul Ehrlich	1056
Optimal Loading of a Thermoelectric Generator	James F. Perkins	1058
Power Output of Thermoelectric Generators	Maria Telkes	1058
Thermal Diffusivity of Metals at High Temperatures	D. Rosenthal and N. E. Friedmann	1059
Multiple Beam Sharpening of Photoelastic Fringes	Daniel Post	1060
Torsional Pendulum of Low Thermal Inertia	C. Ang and C. Wert	1061
Maxwell's Demon Must Remain Incommunicado	Marcel J. E. Golay	1062
Dynamic Mechanical and Electrical Behavior of Polymethyl Methacrylate	David Telfair	1062
Hardening of Pd by Cathodic Hydrogen	Takesi Sugeno and Masamichi Kowaka	1063
Elastic Relaxation and Grain Boundary Diffusion in Cadmium	Y. L. Yousef and R. Kamel	1064
Measurement of Electromagnetic Field in Resonant Cavities by Introduction of Small Metallic Rings	Aldo Gilardini	1064
Centripetal and Centrifugal Electromagnetophoresis	Alexander Kolin	1065
On Shock Propagation in Brass	H. Becker	1066
On the Conduction of Heat into a Growing Vapor Bubble	H. Kurt Forster	1067
Index to Advertisers		xxii

ROBERT L. SPROULL
Editor

ELMER HUTCHISSON,
EARLE C. GREGG,
Retiring Editors

Board of Editors
MAURICE A. BIOT
RAY E. BOLZ
WALLACE J. ECKERT
ROBERT D. HEIDENREICH
LOUIS A. PIPES
ROMAN SMOLUCHOWSKI
W. R. SMYTHE
ARTHUR V. TOBOLSKY
ALVIN M. WEINBERG

Publication of the
American Institute of Physics

OFFICERS
FREDERICK SEITZ,
Chairman
GEORGE B. PEGRAM,
Treasurer

GOVERNING BOARD
ROBERT F. BACHER
J. W. BUCHTA
KARL K. DARROW
S. A. GOUDSMIT
DEANE B. JUDD
HUGH S. KNOWLES
C. C. LAURITSEN
W. H. MARKWOOD, JR.
WILLIAM F. MEGGERS
PHILIP M. MORSE
BRIAN O'BRIEN
HARRY F. OLSON
GEORGE B. PEGRAM
ERIC RODGERS
RALPH A. SAWYER
FREDERICK SEITZ
WILLIAM SHOCKLEY
RICHARD M. SUTTON
J. H. VAN VLECK
MARK W. ZEMANSKY

Consultant on Printing
MELVIN LOOS

ADMINISTRATIVE STAFF
HENRY A. BARTON,
Director
WALLACE WATERFALL,
Executive Secretary
THEODORE VORBURGER,
Advertising Manager
RUTH F. BRYANS,
Publication Manager
EDITH I. NEFTEL
Circulation Manager
KATHRYN SETZE
Assistant Treasurer

The JOURNAL OF APPLIED PHYSICS, published monthly at Prince and Lemon Sts., Lancaster, Pa., is devoted to physics in its role as the science basic to other natural sciences and to the arts and industries. Previous to 1937 this publication was known as PHYSICS. The JOURNAL OF APPLIED PHYSICS publishes editorials and reviews, as well as technical papers of applied physics. Articles appearing in it are indexed regularly, according to their field, by Physics Abstracts, Chemical Abstracts, Engineering Index, Inc., Industrial Arts Index, etc.

Manuscripts should be submitted to Robert L. Sproull, Editor, Rockefeller Hall, Cornell University, Ithaca, New York. Unless otherwise stated, submission of a manuscript is a representation that it has been neither copyrighted, published, nor submitted for publication elsewhere. The authors' institutions are requested to pay a **publication charge** of **$8** per page which, if honored, entitles them to 100 reprints without further charge. Instructions will be sent with galley proofs.

Subscriptions, renewals, and orders for back numbers should be addressed to the American Institute of Physics, 57 East 55 Street, New York 22, New York.

Subscription Price	U.S. and Canada	Elsewhere
To members of the American Institute of Physics and Affiliated Societies	$10.00	$11.00
To all others	12.00	13.00

Back Number Prices

Yearly back number rate when complete year is available: **$14.00.** Single copies: **$1.25** each.

Changes of address should be sent to the American Institute of Physics; **corrected proofs** should be sent to the Publication Manager, American Institute of Physics. **Advertising rates** supplied on request. Orders, advertising copy, and cuts should be sent to the American Institute of Physics.

Entered as second class matter January 22, 1937, at the Post Office at Lancaster, Pennsylvania, under the Act of March 3, 1879. Accepted for mailing at the special rate of postage provided for in paragraph (d-2), section 34.40, P. L. & R. of 1948, authorized May 2, 1932.

Journal
of
Applied Physics

Volume 25, Number 8 August, 1954

A Theorem Concerning Noise in Electron Streams

J. R. PIERCE
Bell Telephone Laboratories, Murray Hill, New Jersey
(Received October 26, 1953)

It is proposed that if I_{max} and I_{min} are the maximum and minimum noise currents in a noise standing wave on an electron stream, then the product of their magnitudes cannot be smaller than

$$|I_{max}||I_{min}|=\alpha eI_0B\left(\frac{\omega}{\omega_q}\right)\left(\frac{kT_c}{eV_0}\right).$$

Various assumptions lead to a factor α ranging between 0.8 and 1. I_0 is beam current, B is band width, ω is frequency, ω_q is effective plasma frequency, T_c is cathode temperature, and V_0 is the voltage specifying the velocity of the electrons.

IT is proposed that if I_{max} and I_{min} are the maximum and minimum noise currents in a noise standing wave on an electron stream, then the product of these cannot be smaller than

$$|I_{max}||I_{min}|=\alpha eI_0B\frac{\omega}{\omega_q}\frac{kT_c}{eV_0}.$$

Here ω_q is effective plasma frequency, I_0 is beam current, V_0 is beam voltage, and T_c is cathode temperature and α is a factor near to unity. Of course, the noise might be higher than this for reasons not considered here.

Here linearity is assumed throughout and it is assumed that the noise current on the electron stream can be represented by a pair of unattenuated waves. Suppose two unattenuated modes of propagation are coupled by a lossless passive transducer. Suppose the modes carry power in opposite directions. Then if P_a and Q_a are the amplitudes of the modes at the input end, and if P_d and Q_d are the amplitudes of the modes at the output end,[1]

$$P_d=(1+k^2)^{\frac{1}{2}}e^{j\theta_1}P_a+ke^{j\theta_2}Q_a \tag{1}$$

$$Q_d=(1+k^2)^{\frac{1}{2}}e^{j\theta_2}Q_a+ke^{j(\theta_1-\theta_2+\theta_3)}P_a. \tag{2}$$

This is the most general possible relation between P_d, Q_d and P_a, Q_a allowed by the conservation of energy.

[1] J. R. Pierce, J. Appl. Phys. **25**, 179 (1954).

Here it has been assumed that the power flows are given by

$$P_aP_a^*,\quad -Q_aQ_a^*$$
$$P_dP_d^*,\quad -Q_dQ_d^*.$$

If we have instead power flows given by

$$K_aP_aP_a^*,\quad -K_aQ_aQ_a^*$$
$$K_dP_dP_d^*,\quad -K_dQ_dQ_d^*,$$

our equations become

$$P_d(K_d)^{\frac{1}{2}}=(1+k^2)^{\frac{1}{2}}(K_a)^{\frac{1}{2}}e^{j\theta_1}P_a+k(K_a)^{\frac{1}{2}}e^{j\theta_2}Q_a \tag{3}$$

$$Q_d(K_d)^{\frac{1}{2}}=(1+k^2)^{\frac{1}{2}}(K_a)^{\frac{1}{2}}e^{j\theta_2}Q_a+k(K_a)^{\frac{1}{2}}e^{j(\theta_1-\theta_2+\theta_3)}P_a. \tag{4}$$

Let us now apply these relations to the case of space-charge waves, in the case of which P_a, P_d, Q_a and Q_d may be electron convection currents. It is assumed that the phase constants of the two waves can be written $\beta=(\omega\mp\omega_q)/u_0$. We can define an impedance K and a voltage V which exhibit the proper behavior at discontinuities in ω_q, the effective plasma frequency, and V_0, the beam voltage with respect to the cathode, and which satisfy the usual relations among voltage, current and power,

$$K=\pm 2\frac{\omega_q}{\omega}\frac{V_0}{I_0} \tag{5}$$

$$V=-\frac{u_0}{e/m}v. \tag{6}$$

Here v is ac electron velocity and u_0 is average electron velocity. The upper sign applies for the fast space charge were and the lower sign for the slow space-charge wave.

Let us consider the case of a standing wave in which

$$Q_{a1}=P_{a1}=P_1/2. \tag{7}$$

This corresponds to maximum current and zero voltage at the input. We have

$$P_{d1}=\left(\frac{K_a}{K_d}\right)^{\frac{1}{2}}e^{j\theta_1}[(1+k^2)^{\frac{1}{2}}+ke^{j(\theta_2-\theta_1)}]P_{a1} \tag{8}$$

$$Q_{d1}=\left(\frac{K_a}{K_d}\right)^{\frac{1}{2}}e^{j\theta_2}[(1+k^2)^{\frac{1}{2}}+ke^{-j(\theta_2-\theta_1)}]P_{a1}. \tag{9}$$

We see that

$$|P_{d1}^2|=|Q_{d1}^2|. \tag{10}$$

That is, if there is a standing wave at the input there is a standing wave at the output. This is necessary, because a standing wave means no power flow, and if there is no power flow at the input there can be no power flow at the output.

Let us now consider another standing wave at the input, for which

$$Q_{a2}=-P_{a2}=-P_{II}/2. \tag{11}$$

This corresponds to zero current and maximum voltage at the input. We have

$$P_{d2}=\left(\frac{K_a}{K_d}\right)^{\frac{1}{2}}e^{j\theta_1}[(1+k^2)^{\frac{1}{2}}-ke^{j(\theta_2-\theta_1)}]P_{a2} \tag{12}$$

$$Q_{d2}=-\left(\frac{K_a}{K_d}\right)^{\frac{1}{2}}e^{j\theta_2}[(1+k^2)^{\frac{1}{2}}-ke^{-j(\theta_2-\theta_1)}]P_{a2}. \tag{13}$$

As before, we have

$$|P_{d2}|=|Q_{d2}|. \tag{14}$$

Suppose that P_I and P_{II} [see (7) and (11)] represent uncorrelated narrow-band noises. Then the resulting signals on the beam following the transducer will be of the form

$$P_{d1}+Q_{d1}=I_1\cos[\tfrac{1}{2}(\beta_p-\beta_q)z-\varphi_1] \tag{15}$$

$$P_{d2}+Q_{d2}=I_2\cos[\tfrac{1}{2}(\beta_p-\beta_q)z-\varphi_2]. \tag{16}$$

Here β_p and β_q are the phase constants of the P and Q modes following the transducer. At any z position the mean square amplitude will be the sum of the squares of the two signals of (15) and (16). It is easy to show that if I_{max}^2 and I_{min}^2 are the maximum and minimum values of the sum of the squares of (15) and (16), then

$$I_{max}^2I_{min}^2=I_1^2I_2^2\sin^2(\varphi_2-\varphi_1). \tag{17}$$

From (8), (9), (12), and (13), I_1, I_2, φ_1, φ_2 can be evaluated and after a great deal of repulsive algebra we find

$$I_{max}^2I_{min}^2=\left(\frac{K_a}{K_d}\right)^2P_I^2P_{II}^2 \tag{18}$$

or

$$|I_{max}||I_{min}|=\frac{K_a}{K_d}|P_I||P_{II}|. \tag{19}$$

We should note that with a suitable transducer it is possible to make $|I_{max}|=|I_{min}|$. This can be done through the use of space-charge-wave amplification and deamplification. Thus, if there is no noise standing wave

$$|I_{max}|^2=|I_{min}|^2=\frac{K_a}{K_d}|P_I||P_{II}|. \tag{20}$$

Equations (19) and (20) seem to indicate that what might be called the "noisiness" of an electron beam depends entirely on the source of the electrons and not on the particular way in which the electrons are accelerated; that is, for a beam of a given current, voltage and geometry, we have a choice only as to the noise standing-wave ratio, but $|I_{max}||I_{min}|$ will be a constant as long as the source of electrons is constant.

Suppose that this is so. Consider a beam of electrons from a thermionic cathode at temperature T_c. We can compute $|I_{max}|^2$ and $|I_{min}|^2$ for the case of electrons accelerated quickly from a temperature-limited cathode. We obtain[2]

$$|I_{max}|^2=2eI_0B \tag{21}$$

$$|\bar{I}_{min}|^2=2eI_0B\left(\frac{\omega}{\omega_q}\right)^2\frac{1}{4}\left(\frac{kT}{eV_0}\right)^2. \tag{22}$$

Accordingly, we would expect in any case

$$\frac{|I_{max}||I_{min}|}{2eI_0B}=\frac{1}{2}\left(\frac{\omega}{\omega_q}\right)\left(\frac{kT}{eV_0}\right). \tag{23}$$

Matters are not quite so simple, however, for we are making some sort of error in dealing with multi-velocity flow. For instance, consider matters at the cathode. There the fluctuations in current and velocity, I_c and v_c, are such that

$$I_c^2=2eI_0B \tag{24}$$

$$v_c^2=(4-\pi)\left(\frac{e}{I_0}\right)\left(\frac{kT_c}{m}\right)B. \tag{25}$$

Let subscripts 1 refer to cathode and quantities without subscripts 1 refer to beam. We have the single-valued-velocity relation of current to velocity

$$I=\frac{\omega}{\omega_{q1}}\frac{I_0}{u_{01}}v, \tag{26}$$

[2] J. R. Pierce, Proc. Inst. Radio Engrs. **40**, 1675–1680 (1952).

and

$$K_a=2\frac{\omega_{q1}}{\omega}\frac{V_{01}}{I_0} \tag{27}$$

$$K_d=2\frac{\omega_q}{\omega}\frac{V_0}{I_0}. \tag{28}$$

From (24)–(28) we conclude that

$$\frac{|I_{\max}I_{\min}|}{2eI_0B}=\frac{1}{2}\left(\frac{\omega}{\omega_q}\right)\left[\frac{2(4-\pi)kT_c}{u_{01}^2m}\right]^{\frac{1}{2}}\left(\frac{V_{01}}{V_0}\right). \tag{29}$$

How are we to interpret V_{01} and u_{01}? We have the single-valued-velocity relation

$$u_{01}^2=2\frac{e}{m}V_{01}. \tag{30}$$

We have for the electrons leaving the cathode

$$\bar{u}^2=\frac{\pi kT_c}{2m} \tag{31}$$

$$\langle u^2\rangle_{Av}=\frac{2kT_c}{m}. \tag{32}$$

If we take

$$u_{01}^2=2\frac{e}{m}V_{01}=\bar{u}^2=\frac{\pi kT_c}{2m}, \tag{33}$$

then we obtain[3]

$$\frac{|I_{\max}I_{\min}|}{2eI_0B}=(0.822)\left(\frac{1}{2}\frac{\omega}{\omega_q}\frac{kT_e}{eV_0}\right). \tag{34}$$

If we take

$$u_{01}^2=2\frac{e}{m}V_{01}=\langle u^2\rangle_{Av}=\frac{2kT_c}{m}, \tag{35}$$

then we obtain

$$\frac{|I_{\max}||I_{\min}|}{2eI_0B}=0.926\left(\frac{1}{2}\frac{\omega}{\omega_q}\frac{kT_e}{eV_0}\right). \tag{36}$$

These examples merely show the magnitude of the error that may be involved, but do not exhaust possible choices of averages. We may say

$$\frac{|I_{\max}||I_{\min}|}{2eI_0B}=\frac{\alpha}{2}\frac{\omega}{\omega_q}\frac{kT_e}{eV_0}, \tag{37}$$

where α is a factor near to unity.

It is interesting to consider also the maxima and minima of the ac velocity fluctuations, $v_{\max}$ and $v_{\min}$. For a standing wave, the ratio of v to I is given by (5) and (6); the velocity corresponding to the current maximum is found at the current minimum and vice versa. From (5), (6), (25), and (29) we obtain

$$\frac{|v_{\max}||v_{\min}|}{v_c^2}=\left(\frac{\omega_q}{\omega}\right)\left(\frac{2eV_{01}}{(4-\pi)kT_c}\right)\times\left(\frac{2(4-\pi)kT_c}{u_{01}^2m}\right)^{\frac{1}{2}}. \tag{38}$$

Using (29) and (36) we can write this

$$\frac{|v_{\max}||v_{\min}|}{v_c^2}=\frac{2\alpha}{(4-\pi)}\frac{\omega_q}{\omega}=2.33\alpha\frac{\omega_q}{\omega}. \tag{39}$$

Either slightly different noise currents and velocities should be obtained in different physical situations, or else there is some proper way of using averages which makes α unity.

Strictly, the foregoing relations apply only when there is not enough space charge to cause a potential minimum in front of a cathode. They will apply in the case of a potential minimum only if the noise in the convection current passing the potential minimum is shot noise. The writer is inclined to believe that in the centimeter range this is the case, and that the relations apply for space-charge-limited flow.

[3] This same factor can be obtained following Dean A. Watkins Noise Reduction in Beam Type Amplifiers, Electronics Research Laboratory, Stanford University, California, Technical Report No. 31; and by using Llewellyn and Peterson's equations to relate noise velocity and current at the potential minimum to noise in the beam.

The Transfer Function as a Tool in the Analysis of a Resistance-Capacitance Coupled Voltage Amplifier

D. L. MING AND R. W. THORPE
Aerojet Engineering Corporation, Azusa, California

(Received July 13, 1953)

The frequency response of a typical resistance-capacitance coupled voltage amplifier has been theoretically analyzed and the result compared with that given by Terman. The advantages and disadvantages of the two methods have been illustrated both mathematically and graphically. A solution to the phase-shift problem at any frequency considered has been included.

I. INTRODUCTION

A THEORETICAL analysis was made on the frequency response of a typical resistance-capacitance coupled amplifier and the result was compared with that of the most common analytical method advanced by Terman.[1] For an amplifier of this type, the response curve rises and falls at the same rate for the respective low- and high-frequency signals.

In Terman's method, the total equivalent circuit of the system breaks down into three separate circuits: A high-band, a mid-band, and a low-band circuit. The gain points of the system are computed from the applicable equivalent circuits, and the over-all response curve is plotted using a set of universal curves. Although this method is based on several approximations, it has been shown to be valid for amplifiers of standard design. For an uncommon design, involving extremely high- or low-value components or unusual ratios between component values, Terman's method may not hold. Moreover, it offers no convenient way of determining the phase shift, since the mid-band equivalent circuit completely eliminates all reactances, and the 180° phase shift indicated occurs only at one particular frequency.

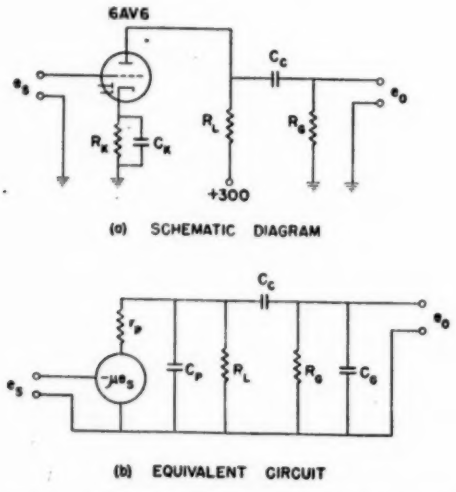


FIG. 1. Circuit diagram of a typical resistance-capacitance coupled voltage amplifier.

[1] F. E. Terman, *Radio Engineering* (McGraw-Hill Book Company, Inc., 1947), third edition.

The theoretical treatment herein presents a method which provides an accurate solution at any frequency considered, and gives complete phase-shift information. However, this method is not intended to replace other methods now in use. It is offered merely as an additional tool for precise determination of phase shift or accurate calculation of gain.

II. DISCUSSION

Based on the present analytical method, the schematic circuit diagram of a typical resistance-capacitance coupled voltage amplifier using the triode section of a 6AV6 receiving tube is presented in Fig. 1 (a), with the total equivalent circuit shown in (b). It will be noted in (b) that both the bias voltages and the power supply voltages are eliminated in this system, but the signal circuit is retained. The values of the circuit components, however, are typical for this tube in either case.

The voltage, current, and impedance relationships, represented in matrix form and properly applied, give

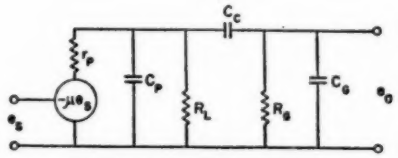


FIG. 2. Total equivalent circuit.

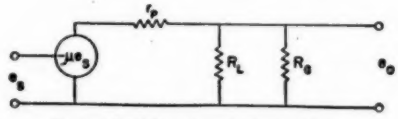


FIG. 3. Mid-band equivalent circuit.

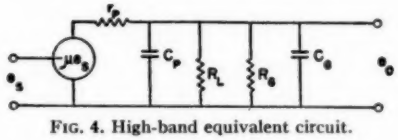


FIG. 4. High-band equivalent circuit.

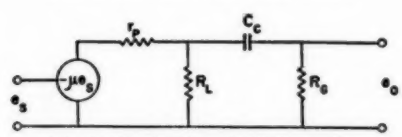


FIG. 5. Low-band equivalent circuit.

FIGS. 2–5. Terman's method as applied to the same circuit.

rise to an expression for the gain of the stage, expressed as a function of frequency. This function will be of the form

$$\frac{e_0}{e_s}=\frac{-AS}{(T_1S+1)(T_2S+1)}, \tag{1}$$

where the quantities are defined under Definition of Symbols.

Equation (1) yields complete information on amplitude and phase with respect to frequency. In using the quadratic formula for finding the factors of the denominator, (T_1S+1) and (T_2S+1), extreme care must be exercised to avoid the introduction of appreciable errors. In the case of the present example, it was found preferable to use a trial-and-error method, as shown in the calculations. Since slide-rule accuracy is not sufficient, logarithms were used in computing the answer.

For purposes of comparison, the same circuit was analyzed using Terman's method, and the results, expressed in db, were plotted to the same graphical scales as before (Figs. 2–5). The two gain curves differed by a maximum of one db, with the -3 db points occurring at substantially the same frequencies. Phase information, of course, was not available from Terman's method. The equivalent circuit shown in Fig. 1(b) may be represented as that given in Fig. 6.

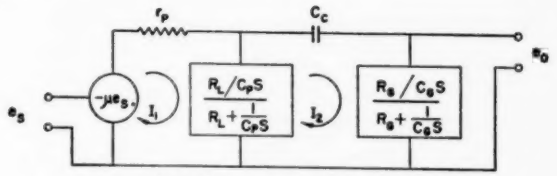


FIG. 6. Modified total equivalent circuit.

Let

$$T_1=R_LC_P$$

and

$$T_2=R_GC_G$$

then,

$$\frac{R_L/C_PS}{R_L+(1/C_PS)}=\frac{R_L}{(T_1S+1)} \tag{2}$$

and

$$\frac{R_G/C_GS}{R_G+(1/C_GS)}=\frac{R_G}{(T_2S+1)}. \tag{3}$$

Then, in matrix form,

$$\begin{vmatrix} r_p+\frac{R_L}{T_1S+1} & -\frac{R_L}{T_1S+1} \\ -\frac{R_L}{T_1S+1} & \frac{1}{C_cS}+\frac{R_L}{T_1S+1}+\frac{R_G}{T_2S+1} \end{vmatrix} \times \begin{vmatrix} I_1 \\ I_2 \end{vmatrix} = \begin{vmatrix} -\mu e_s \\ 0 \end{vmatrix}, \tag{4}$$

where

$$I_2=\frac{-\mu e_s\left(\frac{R_L}{T_1S+1}\right)}{\left[r_p+\frac{R_L}{T_1S+1}\right]\left[\frac{1}{C_cS}+\frac{R_L}{T_1S+1}+\frac{R_G}{T_2S+1}\right]-\left[\frac{R_L}{T_1S+1}\right]^2},$$

$$e_0=I_2\frac{R_G}{T_2S+1}.$$

Then

$$\frac{e_0}{e_s}=\frac{-\mu\frac{R_LR_G}{(T_1S+1)(T_2S+1)}}{\frac{r_p}{C_cS}+\frac{r_pR_L}{T_1S+1}+\frac{r_pR_G}{T_2S+1}+\frac{R_L}{C_cS(T_1S+1)}+\frac{R_LR_G}{(T_1S+1)(T_2S+1)}} \tag{5}$$

and, expanding, rearranging, and collecting the terms

$$\frac{e_0}{e_s}=\frac{-\mu\frac{R_LR_G}{R_L+r_p}C_cS}{S^2\left[\frac{r_pT_1T_2}{r_p+R_L}+\frac{r_pR_LC_cT_2}{r_p+R_L}+\frac{r_pR_GC_cT_1}{r_p+R_L}\right]+S\left[\frac{r_pT_1}{r_p+R_L}+\frac{r_pT_2}{r_pR_L}+\frac{r_pR_LC_c}{r_p+R_L}+\frac{r_pR_GC_c}{r_p+R_L}+\frac{R_LT_2}{r_p+R_L}+\frac{R_LR_GC_c}{r_p+R_L}\right]+1}. \tag{6}$$

Assign the following values:

$$T_1=R_LC_P,\quad T_2=R_GC_G,\quad T_3=R_LC_c,\quad T_4=R_GC_c,\quad K_1=R_L,$$
$$K_2=\frac{1}{r_p+R_L},\quad K_3=r_p,\quad K_4=K_1K_2,\quad K_5=K_2K_3,\quad (K_4+K_5=1).$$

By substitution:

$$\frac{e_0}{e_s}=\frac{-\mu K_4T_4S}{S^2[K_5(T_1T_2+T_2T_3+T_1T_4)]+S[T_2+T_4+K_5(T_1+T_3)]+1}. \tag{7}$$

Equation (7) is the general expression for the transfer function of the circuit shown in Fig. 6.

For the example presented here, the following values were chosen for the components:

$$\mu=100,\quad C_c=0.006\times10^{-6},$$
$$r_p=0.11\times10^6,\quad R_G=0.47\times10^6,$$
$$C_P=0.8\times10^{-12},\quad C_G=C_{GK}+C_{GP}(1+A)=140\times10^{-12},$$
$$R_L=0.47\times10^6,$$

where the resistances are given in ohms and the capacitances in farads. Then, numerically the quantities are

$$T_1=0.376\times10^{-6},\quad K_2=1.722\times10^{-6},$$
$$T_2=65.9\times10^{-6},\quad K_3=0.11\times10^6,$$
$$T_3=2820\times10^{-6},\quad K_4=0.8103,$$
$$T_4=2820\times10^{-6},\quad K_5=0.1897.$$
$$K_1=0.47\times10^6,$$

Substituting these values in Eq. (7), one obtains

$$\frac{e_0}{e_s}=\frac{-0.228S}{S^2[0.03538\times10^{-6}]+S[3421\times10^{-6}]+1}.$$

This expression is of the form

$$\frac{e_0}{e_s}=\frac{-AS}{BCS^2+(B+C)S+1}=\frac{-AS}{(BS+1)(CS+1)} \tag{8}$$

where $BC=0.03538\times10^{-6}=35\,380\times10^{-12}$ and

$$B+C=3421\times10^{-6}.$$

A trial-and-error method of simultaneous solution using logarithms yields

$$B=3411\times10^{-6}$$

and

$$C=10.38\times10^{-6},$$

$$\frac{e_0}{e_s}=\frac{-0.228S}{(3411\times10^{-6}S+1)(10.38\times10^{-6}S+1)}=\frac{-0.228S}{\left(\frac{S}{293}+1\right)\left(\frac{S}{96\,340}+1\right)}.$$

This latter expression may be represented as a single vector, having an absolute amplitude $|R|$, and a phase angle of ϕ. The amplitude expressed in db may be found as shown in Eq. (9).

$$|R|(\text{db})=\underbrace{20\text{Log }0.228}_{1}+\underbrace{20\text{Log}\omega}_{2}+\underbrace{20\text{Log}\frac{1}{\sqrt{\left(\frac{\omega}{293}\right)^2+1}}}_{3}+\underbrace{20\text{Log}\frac{1}{\sqrt{\left(\frac{\omega}{96,340}\right)^2+1}}}_{4}, \tag{9}$$

where

(1) $=20\log(0.228)=-13.84$,

(2) $=20\log\omega$,
slope $=+20$ db/decade, through $R=0$, $\omega=1$,

(3) $=0$, when $\omega\ll1$,
$=20\log(1/\omega)$, when $\omega\gg1$.

The breakpoint occurs at 293 radians/second, and the slope is equal to -20 db/decade.

(4) $=0$, when $\omega\ll1$,
$=20\log(1/\omega)$, when $\omega\gg1$.

The breakpoint occurs at 96 340 radians/second, and the slope is equal to -20 db/decade.

From this information, the curve of gain ($|R|$) *vs* frequency may be plotted as shown in Fig. 7.

The phase shift may be determined as shown in Eq. (10). Thus,

$$\phi=\underbrace{-90^\circ}_{\phi_1}\underbrace{-\text{Tan}^{-1}(3411\times10^{-6}\omega)}_{\phi_2}\underbrace{-\text{Tan}^{-1}(10.38\times10^{-6}\omega)}_{\phi_3} \tag{10}$$

where the values of the quantities are as tabulated in Table I. From this information, the curve of the phase shift *vs* frequency may be plotted as shown in Fig. 7.

$$\text{Mid-band gain}=\mu\frac{(R_GR_L/R_G+R_L)}{r_p+\frac{R_GR_L}{R_G+R_L}}$$
$$=68.1.$$

Expressed in db, gain $=36.7$.

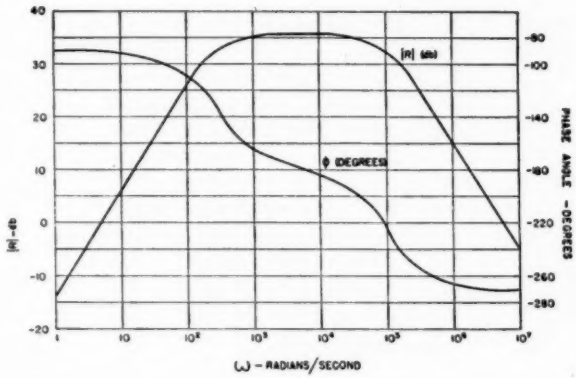


FIG. 7. Frequency response curve from transfer function method

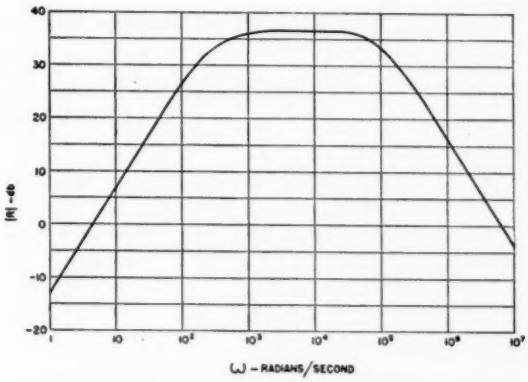


FIG. 8. Frequency response curve from Terman's method.

The low-frequency half-power point occurs when

$$X_{c_c}=R_G+\frac{R_L r_p}{R_L+r_p}=0.5591\times 10^6$$

which is at frequency

$$f_1=\frac{10^6}{2\pi(0.006)(0.5591\times 10^6)}=47.5 \text{ cps.}$$

$$=298 \text{ radians/second.}$$

The high-frequency half-power point occurs when

$$X_{c_g}=\frac{R_G R_L r_p}{R_G R_L+R_L r_p+r_p R_G}=0.075\times 10^6$$

which is at frequency

$$f_2=\frac{10^{12}}{(0.075\times 10^6)2\pi(140)}=15\,170 \text{ cps}$$

$$=95\,400 \text{ radians/second.}$$

TABLE I.

ω	ϕ_1	ϕ_2	ϕ_3	ϕ
1	−90	0	0	− 90
10	↓	2	0	− 92
10^2		19	0	−109
10^3		73	1	−164
2×10^3		82	1	−173
10^4		88	6	−184
3×10^4		89	17	−196
10^5		90	45	−225
10^6		90	85	−265
10^7	↓	90	89	−269

The two half-power (−3 db) points and the mid-band gain being known, Terman's universal curves can be used to plot the frequency response as shown in Fig. 8.

DEFINITIONS OF SYMBOLS

C_c = Blocking capacitor.
C_{GK} = Capacitance between the grid and cathode of following stage.
C_{GP} = Capacitance between the grid and plate of following stage.
C_K = Cathode bypass capacitor.
C_P = Output capacity of tube.
I_1 = Loop current as indicated in Fig. 6.
I_2 = Loop current as indicated in Fig. 6.
K_1 = A constant, $=R_L$,
K_2 = A constant, $=1/(r_p+R_L)$.
K_3 = A constant, $=r_p$.
K_4 = A constant, $=K_1K_2=R_L/(r_p+R_L)$.
K_5 = A constant, $=K_2K_3=r_p/(r_p+R_L)$.
$|R|$ = The absolute amplitude of the transfer function, without regard to phase angle.
R_G = Grid leak of following stage.
R_K = Cathode bias resistor.
R_L = Plate load resistor.
S = Operator $(j\omega)$.
T_1 = Time constant of R_L and C_p.
T_2 = Time constant of R_G and C_G.
T_3 = Time constant of C_c and R_L.
T_4 = Time constant of C_c and R_g.
e_0 = Output voltage.
e_g = Signal (input) voltage.
f_1 = Low frequency half-power point.
f_2 = High frequency half-power point.
r_p = ac Plate resistance of tube.
μ = Amplification factor of tube.
ϕ = Phase angle,
ω = Frequency in radians per second.

Geometrical Acoustics. I. The Theory of Weak Shock Waves*

JOSEPH B. KELLER
Institute of Mathematical Sciences, New York University,† New York, New York
(Received April 24, 1953)

In the first part the discontinuity conditions in an arbitrary continuous material are deduced for a general (i.e., possibly curved) discontinuity surface. It is then shown that only three types of discontinuities are possible—shocks, contact discontinuities, and phase-change fronts. In the second part the acoustic discontinuity conditions are deduced and specialized to a perfect fluid without heat conduction. Then a first-order partial differential equation is obtained for the location of an acoustic shock front. This equation can be solved, as in optics, by means of rays. The variation of shock strength along a ray is then determined (this is one main result of this paper). Coefficients of reflection and transmission for an acoustic shock at a contact discontinuity in the basic flow are also obtained. Finally, the results are exemplified by an analysis of the shock tube.

I. INTRODUCTION

ONE of the main difficulties in the theoretical analysis of shock wave problems is that the shock front and the flow behind it must be determined simultaneously, since they affect each other. However, for very weak (acoustic) shocks the interaction is so slight that the motion of the shock front and the variation in its strength can be determined independently of the rest of the flow. This fact is deduced in the present report, and the main consequences of it are also determined. The result is that weak shocks can be analyzed by the methods of geometrical optics, and for this reason the theory of weak shocks may be called geometrical acoustics. Another acoustical phenomenon —the propagation of periodic sound waves of high frequency—can also be analyzed by the methods of geometrical acoustics, and this analysis will be presented in another report.

Although it has often been noticed that certain shock wave phenomena could be described in optical terms, there does not seem to be any complete derivation of geometrical acoustics in the literature. The present derivation begins with a deduction of the discontinuity conditions for curved discontinuities in any continuous medium, and continues with the analysis of the possible discontinuities in arbitrary media. Then the acoustic equations are deduced as the variational equations of the nonlinear fluid dynamic equations. In the same way the acoustic discontinuity conditions are obtained, and it is shown how the geometrical optics, considerations of wave fronts, rays, etc., apply to the determination of shock waves in perfect fluids. Finally, as one of the main results, the variation of shock strength along a ray is determined. This result is extended to the determination of the reflection and transmission coefficients for shocks incident on contact discontinuities (boundaries).

For the most effective use of the present theory it is important to take account of diffracted rays and wavefronts, which can be done by slightly refining the methods of ordinary geometrical optics. An explanation of this refinement will be presented in a future report, although use has already been made of it.[1]

The present analysis is patterned after that of R. K. Luneburg's study of the propagation of discontinuities in electromagnetic theory.[2]

II. DISCONTINUITY CONDITIONS IN CONTINUUM MECHANICS

A. Fundamental Equations

The fundamental equations of continuum mechanics, in their integral form, are

$$\frac{d}{dt}\int_V \rho dV+\int_S \rho(u_i-v_i)dS_i=0, \quad \text{Mass} \tag{1}$$

$$\frac{d}{dt}\int_V \rho u_i dV=\int_V \rho f_i dV + \int_S [T_{ij}-\rho u_i(u_j-v_i)]dS_j, \quad \text{Linear momentum} \tag{2}$$

$$\frac{d}{dt}\int_V \rho(\epsilon+\tfrac{1}{2}u_i^2)dV=\int_V \rho f_i u_i dV + \int_S [T_{ij}u_i-q_j-\rho(\epsilon+\tfrac{1}{2}u_i^2)(u_j-v_j)]dS_j, \quad \text{Energy} \tag{3}$$

$$\frac{d}{dt}\int_V \rho(x_{i+1}u_{i+2}-x_{i+2}u_{i+1})dV = \int_V \rho(x_{i+1}f_{i+2}-x_{i+2}f_{i+1})dV + \int_S [x_{i+1}T_{i+2,j}-x_{i+2}T_{i+1,j} -\rho(x_{i+1}u_{i+2}-x_{i+2}u_{i+1})(u_j-v_j)]dS_j, \quad \text{Angular momentum} \tag{4}$$

* A revision of notes on lectures on Mechanics of Continuous Media delivered at New York University in the fall of 1949 and at the University of California (Berkeley) in June, 1951.

† This report represents results obtained at the Institute for Mathematics and Mechanics, New York University, under the auspices of Contract No. nr-285(02).

[1] See, for example, Albert Blank and Joseph B. Keller, Comm. Pure Appl. Math., Vol. IV, No. 2 (1951).

[2] See, for example, R. K. Luneburg, "Mathematical theory of optics," Brown University (1944); "Propagation of electromagnetic waves," New York University (1949).

$$\frac{d}{dt}\int_V \rho s dV - \int_S [q_i\theta^{-1}+\rho s(u_i-v_i)]dS_i \geq 0, \quad \text{Second law} \quad (5)$$

$$\epsilon=\epsilon(\rho,s), \quad \text{First law} \quad (6)$$

$$\theta=\epsilon_s. \quad (7)$$

In these equations $S(t)$ is the surface of an arbitrary volume $V(t)$, x_i are Cartesian coordinates, i takes the values 1, 2, 3 (mod 3), ρ is density, u_i is particle velocity, s is entropy per unit volume, ϵ is internal energy, f_i is external force (both per unit mass), T_{ij} is the stress tensor, q_i is the heat flow vector, and θ is temperature. The summation convention is understood and all quantities are functions of x_i, t. If the equation of $S(t)$ is

$$\phi(x_i,t)=0, \quad (8)$$

then the unit normal ν_i has components

$$\nu_i=\phi_{x_i}(\phi_{x_j}{}^2)^{-\frac{1}{2}}. \quad (9)$$

The velocity v_i of the surface is then defined by

$$v_i=-\nu_i\phi_t(\phi_{x_j})^{-\frac{1}{2}}. \quad (10)$$

It is assumed that the normal ν_i points out of V, and that $dS_i=\nu_i dS$ where dS is an element of area of S.

The preceding equations can be transformed into other forms by means of the two transformation theorems

$$\frac{d}{dt}\int_V A dV=\int_V A_t dV+\int_S A v_i dS_i, \quad (11)$$

$$\int_V (B_i)_{x_i} dV=\int_S B_i dS_i. \quad (12)$$

For the validity of the first theorem it is necessary that A be continuous, and we will have to make use of this fact.

Applying (11) to the left side of (1) yields

$$\int_V \rho_t dV+\int_S \rho v_i dS_i+\int_S \rho(u_i-v_i)dS_i=0. \quad (13)$$

Now making use of (12) in (13) we find

$$\int_V [\rho_t+(\rho u_i)_{x_i}]dV=0. \quad (14)$$

Since (14) holds for an arbitrary volume V, we have when the integrand is continuous

$$\rho_t+(\rho u_i)_{x_i}=0. \quad (15)$$

Proceeding similarly with (2)–(5) yields

$$u_{it}+u_j(u_i)_{xj}-\rho^{-1}(T_{ij})_{xj}=f_i, \quad (16)$$

$$\epsilon_t+u_i\epsilon_{x_i}+\rho^{-1}[(q_i)_{x_i}-T_{ij}(u_i)_{x_j}]=0, \quad (17)$$

$$T_{ij}=T_{ji}, \quad (18)$$

$$\rho(s_t+u_i s_{x_i})+(q_i\theta^{-1})_{x_i}\geq 0. \quad (19)$$

Equations (15)–(19) are the differential forms of the fundamental equations of continuum mechanics, and apply only when all the functions ρ, u_i, etc., are continuous.

If any of the functions ρ, u_i, etc., has a jump discontinuity across a surface S', the above differential equations do not apply on S'. In this case we choose a region V which is divided into two parts, V_1 and V_2, by S'. The boundary of V_1 consists of S' and another part which we call S_1; similarly, V_2 is bounded by S' and S_2. Since S' is a discontinuity surface, we indicate the limits which functions attain on approaching S' by a subscript one or two, according as S' is approached through V_1 or V_2. We now apply (1) to each of the regions V, V_1, and V_2 and then subtract the equations for V_1 and V_2 from that for V. We thus obtain

$$-\int_{S'} [\rho(u_i-v_i)]_1 dS_{i1}'-\int_{S'} [\rho(u_i-v_i)]_2 dS_{i2}'=0. \quad (20)$$

Combining these integrals, and noting that $dS_{i1}'=-dS_{i2}'$, we obtain

$$\int_{S'} \{[\rho(u_i-v_i)]i-[\rho(u_i-v_i)]_2\}dS_{i2}'=0. \quad (21)$$

Since S' is an arbitrary part of the discontinuity surface, we have when the integrand in (21) is continuous,

$$[\rho(u_i-v_i)]_1\nu_1=[\rho(u_i-v_i)]_2\nu_i. \quad (22)$$

In eactly the same way, (2), (3), (5) lead to

$$[T_{ij}-\rho u_i(u_j-v_j)]_1\nu_j=[T_{ij}-\rho u_i(u_j-v_j)]_2\nu_j, \quad (23)$$

$$[T_{ij}u_i-q_j-\rho(\epsilon+\tfrac{1}{2}u_i^2)(u_j-v_j)]_1\nu_j = [T_{ij}u_i-q_j-\rho(\epsilon+\tfrac{1}{2}u_i^2)(u_j-v_j)]_2\nu_j, \quad (24)$$

$$[q_i\theta^{-1}+\rho s(u_i-v_i)]_1\nu_i\geq[q_i\theta^{-1}+\rho s(u_i-v_i)]_2\nu_i. \quad (25)$$

In deriving (25) it is assumed that ν_i points toward side one; if not, the inequality must be reversed. The equation obtained in the above manner from (4) is an immediate consequence of (24). Thus, the discontinuity conditions implied by (1)–(5) are (22)–(25). These conditions, together with the differential relations (15)–(19), also imply the integral Eqs. (1)–(5). (The derivation of (25) requires a proof that the left side of (5) computed for V is greater than the sum of the same expressions computed for V_1 and V_2. This follows directly from (19).)

B. Shocks, Contact Discontinuities, and Phase Change Fronts

The jump conditions (22)–(25) may be rewritten in the form

$$\rho_1U_1=\rho_2U_2 \quad \text{or} \quad \rho_2\nu_i\Delta u_i=-U_1\Delta\rho, \quad (26)$$

$$\rho_1U_1\Delta u_i=\nu_j\Delta T_{ij}, \quad (27)$$

$$\rho_1U_1\Delta(\epsilon+\tfrac{1}{2}u_i^2)-\nu_j\Delta(T_{ij}u_i-q_j)=0, \quad (28)$$

$$\rho_1U_1\Delta s+\nu_i\Delta(q_i\theta^{-1})\geq 0. \quad (29)$$

Here we have introduced the normal velocities U_1, U_2 of the surface relative to the medium on sides one and two, respectively, and the "jump" Δ, by

$$\Delta g \equiv g_1 - g_2, \tag{30}$$

$$U_1 = (v_i - u_{i1})\nu_i, \tag{31}$$

$$U_2 = (v_i - u_{i2})\nu_i. \tag{32}$$

If $U_1 = 0$ the surface moves with the medium and the discontinuity is called a contact discontinuity. In this case, since ρ_1 and ρ_2 are not zero, (26)–(29) become

$$U_2 = 0 \quad \text{or} \quad \nu_i \Delta u_i = 0, \tag{33}$$

$$\nu_j \Delta T_{ij} = 0, \tag{34}$$

$$\nu_j \Delta q_j = \nu_j T_{ij} \Delta u_i, \tag{35}$$

$$\nu_i \Delta(q_i \theta^{-1}) \geq 0. \tag{36}$$

Equations (33), (34) show that the normal velocity and the stress on the surface are continuous, while (35) shows that the jump in normal heat flux equals the jump in power of the normal stress.

If $U_1 \neq 0$, (27) may be written

$$\Delta u_i = (\rho_1 U_1)^{-1} \nu_j \Delta T_{ij}. \tag{37}$$

If in addition, $\Delta\rho \neq 0$, the discontinuity is called a shock. From (37), (26) we find U_1, the normal velocity of the shock relative to the medium on side one:

$$U_1 = \pm\left(\frac{-\rho_2 \nu_i \nu_j \Delta T_{ij}}{\rho_1 \Delta\rho}\right)^{\frac{1}{2}}. \tag{38}$$

From (38), (26) we find that U_2 is given by (38) with subscripts one and two interchanged.

If $U_1 \neq 0$ and $\Delta\rho = 0$, the surface is called a phase change front. Equation (37) applies, while (26), (28) yield

$$\nu_i \Delta u_i = 0, \tag{39}$$

$$U_1 = \frac{\nu_j \Delta(u_i T_{ij} - q_j)}{\rho_1 \Delta(\epsilon + \frac{1}{2} u_i^2)}. \tag{40}$$

Equation (39) shows that the normal velocity is continuous, while (40) gives the normal velocity U_1 of the surface relative to the fluid, which is also equal to U_2. Such fronts occur in melting, freezing, evaporation, and condensation, etc.

C. Perfect Fluids

A perfect (or nonviscous) fluid is defined by

$$T_{ij} = -p\delta_{ij}. \tag{41}$$

Here p is the thermodynamic pressure, defined by

$$p = -\frac{\partial \epsilon}{\partial(\rho^{-1})} = \rho^2 \epsilon_\rho. \tag{42}$$

For incompressible media (42) is meaningless, and then p is introduced as a new dependent function.

For a perfect fluid (18) is automatically satisfied while (16), (17), (19) can be reduced to

$$u_{it} + u_j (u_i)_{xj} + \rho^{-1} p_{xi} = f_i, \tag{43}$$

$$s_t + u_i s_{xi} = -\frac{1}{\rho\theta}(q_i)_{xi}, \tag{44}$$

$$q_i \theta_{xi} \leq 0. \tag{45}$$

In deducing (45), the assumption that $\theta > 0$ is made use of.

At a contact discontinuity (34)–(36) become

$$\Delta p = 0, \tag{46}$$

$$\nu_j \Delta q_j = 0, \tag{47}$$

$$q_i \nu_i \Delta(\theta^{-1}) \geq 0. \tag{48}$$

Thus, p and the normal heat flux are continuous, while heat flows from higher to lower temperature at a contact discontinuity.

At a shock (37) and (38) become

$$\Delta u_i = -\nu_i (\rho_1 U_1)^{-1} \Delta p, \tag{49}$$

$$U_1 = \pm\left(\frac{\rho_2 \Delta p}{\rho_1 \Delta\rho}\right)^{\frac{1}{2}}. \tag{50}$$

At a phase change front (49) applies, and this with (39) implies that u_i and p are continuous. Then (40) becomes

$$U_1 = \frac{-\nu_j \Delta q_j}{\rho_1 \Delta\epsilon}. \tag{51}$$

If no heat conduction occurs, then a phase change front cannot occur in a perfect fluid, since (51) would imply $\Delta\epsilon = 0$, and then all quantities would be continuous. (U_1 cannot vanish, by definition of a phase change front.)

III. GEOMETRICAL ACOUSTICS—WEAK SHOCKS

A. The Variational or Acoustic Equations

Let us consider a one-parameter set of solutions of the equations of continuum mechanics depending upon a parameter η. We assume the solutions are differentiable with respect to η at $\eta = 0$ and we differentiate (6), (7), (22)–(24), (42) with respect to η at $\eta = 0$, obtaining

$$\dot\rho_t + (\dot\rho u_i + \rho \dot u_i)_{xi} = 0, \tag{1}$$

$$\dot u_{it} + u_j(\dot u_i)_{xj} + u_j(\dot u_i)_{xj} - \frac{1}{\rho}(\dot T_{ij})_{xj} + \frac{\dot\rho}{\rho^2}(T_{ij})_{xj} = \dot f_i, \tag{2}$$

$$\dot\epsilon_t + \dot u_i \epsilon_{xi} + u_i \dot\epsilon_{xi} + \frac{1}{\rho}[(q_i)_{xi} - \dot T_{ij}(u_i)_{xj} - T_{ij}(\dot u_i)_{xj}] - \frac{\dot\rho}{\rho^2}[(q_i)_{xi} - T_{ij}(u_i)_{xj}] = 0, \tag{3}$$

$$\dot{T}_{ij}=\dot{T}_{ij}, \tag{4}$$

$$\dot{\epsilon}=\epsilon_\rho\dot{\rho}+\epsilon_s\dot{s}, \tag{5}$$

$$\dot{p}=p_\rho\dot{\rho}+p_s\dot{s}, \tag{6}$$

$$\dot{\theta}=\theta_\rho\dot{\rho}+\theta_s\dot{s}. \tag{7}$$

In these equations a dot denotes differentiation with respect to η. These equations are linear in the derivatives $\dot{\rho}$, $\dot{s}$, $\dot{u}_i$, etc., which are called the variations of the corresponding undifferentiated quantities, or the acoustic density, entropy, velocity, etc. The equations are called the variational or acoustic equations, and the undifferentiated quantities are called the basic flow. It is to be noted that a solution of the homogeneous variational equations (with $\dot{f}_i=0$) is given by the x_i derivative of a solution of the original equations (with $f_i=0$).

In order to differentiate the discontinuity conditions, we must take account of the fact that the discontinuity surface may depend upon η. Thus, from II-(8) we have

$$\dot{\phi}+\phi_{x_i}\dot{x}_i=0. \tag{8}$$

Since only normal displacements of the surface need be considered, we may satisfy (8) by setting

$$\dot{x}_i=-\nu_i\dot{\phi}(\phi_{x_j}{}^2)^{-\frac{1}{2}}. \tag{9}$$

Making use of (9), we obtain from (22)–(25)

$$[\rho_2\nu_i\Delta u_i+U_1\Delta\rho]^{\cdot}-[\rho_2\nu_i\Delta u_i+U_i\Delta\rho]_{x_j}\nu_j\dot{\phi}(\phi_{x_k}{}^2)^{-\frac{1}{2}}=0, \tag{10}$$

$$[\rho_1U_1\Delta u_i-\nu_j\Delta T_{ij}]^{\cdot}-[\rho_1U_1\Delta u_i-\nu_j\Delta T_{ij}]_{x_k}\nu_k\dot{\phi}(\phi_{x_k}{}^2)^{-\frac{1}{2}}=0, \tag{11}$$

$$[\rho_1U_1\Delta(\epsilon+\tfrac{1}{2}u_i{}^2)-\Delta(T_{ij}u_i-q_j)\nu_j]^{\cdot}-[\rho_1U_1\Delta(\epsilon+\tfrac{1}{2}u_i{}^2)-\Delta(T_{ij}u_i-q_j)\nu_j]_{x_k}\nu_k\dot{\phi}(\phi_{x_k}{}^2)^{-\frac{1}{2}}=0. \tag{12}$$

The inequalities yield no definite results when differentiated unless the equalities hold for $\eta=0$. Then they also hold for the derivatives, unless only positive η are considered, in which case the inequalities hold for the derivatives. Equations (10)–(12) are linear in the acoustic quantities, and may be called the acoustic discontinuity conditions.

B. Acoustic Discontinuities

When the basic flow is continuous (i.e., $\rho_1=\rho_2$, $U_1=U_2$, etc.), (10)–(12) simplify to

$$\rho\nu_i\Delta\dot{u}_i=-U\Delta\dot{\rho}, \tag{13}$$

$$\rho U\Delta\dot{u}_i=\nu_j\Delta\dot{T}_{ij}, \tag{14}$$

$$\rho U\Delta(\dot{\epsilon}+u_i\cdot\dot{u}_i)=\nu_j\Delta(\dot{T}_{ij}u_i+T_{ij}\dot{u}_i-\dot{q}_j). \tag{15}$$

By using (14), we may simplify (15) to

$$\rho U\Delta\dot{\epsilon}=\nu_j(T_{ij}\Delta\dot{u}_i-\Delta\dot{q}_j). \tag{16}$$

If $U=0$, we have an acoustic contact discontinuity, and (13), (14), and (16) become

$$\nu_i\Delta\dot{u}_i=0, \tag{17}$$

$$\nu_j\Delta\dot{T}_{ij}=0, \tag{18}$$

$$\nu_j\Delta\dot{q}_j=0. \tag{19}$$

If $U\neq0$ but $\Delta\dot{\rho}=0$, we have an acoustic phase change front, and (13) becomes (17), while (14) and (16) become

$$\Delta\dot{u}_i=(\rho U)^{-1}\nu_j\Delta\dot{T}_{ij}, \tag{20}$$

$$U=(\rho\Delta\dot{\epsilon})^{-1}\nu_j(T_{ij}\Delta\dot{u}_i-\Delta\dot{q}_j). \tag{21}$$

If $U\neq0$, $\Delta\dot{\rho}\neq0$, we have an acoustic shock, and (14) becomes (18), which, with (13) yields

$$U=\pm\left(\frac{-\nu_i\nu_j\Delta\dot{T}_{ij}}{\Delta\dot{\rho}}\right)^{\frac{1}{2}}. \tag{22}$$

In a perfect fluid the following further simplifications result:

$$\Delta\dot{p}=0 \qquad \text{contact discontinuity,} \tag{23}$$

$$\Delta\dot{u}_i=-(\rho U)^{-1}\nu_i\Delta\dot{p} \qquad \text{shock or phase front,} \tag{24}$$

$$U=\pm(\Delta\dot{p}/\Delta\dot{\rho})^{\frac{1}{2}} \qquad \text{shock,} \tag{25}$$

$$U=-(\rho\Delta\dot{\epsilon})^{-1}\nu_j\Delta\dot{q}_j \qquad \text{phase front,} \tag{26}$$

$$\rho U\theta\Delta\dot{s}=-\nu_j\Delta\dot{q}_j \qquad \text{any discontinuity.} \tag{27}$$

The last equation follows when II-(7), II-(42), and III-(5) are used in (16).

If we now consider a perfect fluid without heat flow ($q_j=0$), then phase fronts become impossible, and at a shock (27) yields

$$\Delta\dot{s}=0. \tag{28}$$

Using (28) in (5) yields $\Delta\dot{p}=p_\rho\Delta\dot{\rho}$, and thus from (25) we have for the velocity of an acoustic shock

$$U=\pm(p_\rho)^{\frac{1}{2}}=\pm c. \tag{29}$$

The quantity $c=(p_\rho)^{\frac{1}{2}}$ is called the sound speed, and it may vary with position and time since it depends upon the basic flow.

If the definition of U in II-(31) or II-(32) and II-(9), II-(10), are used in (29), we find

$$(\phi_t+u_j\phi_{x_j})^2=c^2(\phi_{x_j}{}^2). \tag{30}$$

This is a first-order partial differential equation which must be satisfied by an acoustic shock in a perfect, nonconducting fluid. The coefficients depend upon the basic flow.

If $\phi_t\neq0$, we may write the equation $\phi=0$ in the form

$$t=W(x_i). \tag{31}$$

Then (30) becomes an equation for W:

$$(u_jW_{x_j}-1)^2=c^2W_{x_j}{}^2. \tag{32}$$

If $u_j=0$, (32) becomes the eiconal equation

$$\frac{1}{c^2}=(W_{x_j})^2=W_{x_1}{}^2+W_{x_2}{}^2+W_{x_3}{}^2. \tag{33}$$

If $u_j=$constant, (32) can be reduced to (33) by making use of the change of variable $\bar{x}_j=x_j-u_jt$. Equations (32) and (33) apply only when c and u_j are independent of t, which we will henceforth assume to be the case.

C. Wave Fronts and Rays

The acoustic shock surfaces $W=$constant will now be called wave fronts because the integration of equations such as (32) has been studied in geometrical optics where that term is used. There is, of course, a complete theory of such equations, and we will merely specialize it to this case.

First we define the wave normal p_i by

$$p_i=W_{x_i}. \tag{34}$$

From (32), (34), p_i satisfies the quadratic equation

$$(u_ip_i-1)^2=c^2p_i{}^2. \tag{35}$$

At each point in space (35) defines a surface containing the endpoints of the possible wave normals p_i. This is called the surface of wave normals. The reciprocal of the radius from the origin to this surface in any direction is the velocity v in that direction of a possible wavefront, since by I-(10), I-(31), I-(32)

$$v=(p_i{}^2)^{-\frac{1}{2}}. \tag{36}$$

By examining (35) one finds that the wave normal surface is a surface of revolution about the direction u_i. Its section by a plane containing this direction is a hyperbola, parabola, or ellipse, according as the basic flow is supersonic, sonic, or subsonic (i.e., according as $u=(u_i{}^2)^{\frac{1}{2}}\gtrless c$). If $u_i=0$, the wave normal surface is a sphere of radius c^{-1}. If θ denotes the angle between p_i and u_i, then in the subsonic case there is one wave normal for each value of θ, and the wave-front velocity $v=c+u\cos\theta$. The same applies to the sonic case, except that there is no wave normal for $\theta=\pi$.

In the supersonic case there are two solutions p_i for each direction if $0\leq\theta<\cos^{-1}(c/u)$, and then $v=u\cos\theta\pm c$, the upper sign applying to that branch of the surface nearer the origin. When $\cos^{-1}(c/u)\leq\theta<\cos^{-1}(-c/u)$, there is one solution p_i for each direction, and $v=c+u\cos\theta$. If $\cos^{-1}(c/u)\leq\theta$ there is no solution. These results are readily understood in terms of vectorial addition of u_i and v_i.

The Hamiltonian function $H(x_i,p_i)$ is defined by

$$H(x_i,p_i)=c^2p_j{}^2-(1-u_jp_j)^2. \tag{37}$$

In terms of this, certain curves $x_i(\sigma)$, which we call rays, are introduced by the equations

$$\frac{dx_i}{d\sigma}=\lambda H_{p_i}=2\lambda[c^2p_i+u_ic(p_j{}^2)^{\frac{1}{2}}]. \tag{38}$$

Here λ is an arbitrary nonzero factor. From the last expression in (38) we see that the rays are orthogonal trajectories of the wave fronts if and only if u_i is zero or parallel to p_i, since then $dx_i/d\sigma$ is parallel to p_i, which is a wave normal.

From the general theory of first-order equations it follows that the rays satisfy Hamilton's equations, Lagrange's equations, Fermat's principle, etc. Furthermore, by means of the rays we can construct the solution of (32) corresponding to any given initial data. The wave fronts can also be found by Huygen's construction. The vector $dx_i/d\sigma$ also lies on Fresnel's ray surface, which is reciprocal to the wave normal surface.

D. Variation of Shock Strength Along a Ray

To determine the variation of shock strength (i.e., $\Delta\dot{\rho}$, $\Delta\dot{p}$, etc.) along a ray, we must consider the acoustic equations for a perfect, nonconducting fluid. By specializing (1)–(3), (6), or by differentiating I-(43), I-(44), we obtain

$$\dot{\rho}_t+(\dot{\rho}u_i+\rho\dot{u}_i)_{x_i}=0, \tag{39}$$

$$\dot{u}_{i_t}+\dot{u}_j(u_i)_{x_j}+u_j(\dot{u}_i)_{x_j}+\rho^{-1}\dot{p}_{x_i}-\dot{\rho}\rho^{-2}p_{x_i}=\dot{f}_i, \tag{40}$$

$$\dot{s}_t+\dot{u}_js_{x_j}+u_j\dot{s}_{x_j}=0, \tag{41}$$

$$\dot{p}=p_s\dot{s}+p_\rho\dot{\rho}. \tag{42}$$

We now consider the quantities $\dot{u}_i(x_j,W(x_j))=\bar{u}_i(x_j)$, $\dot{\rho}(x_j,W(x_j))=\bar{\rho}(x_j)$, $\dot{s}(x_j,W(x_j))=\bar{s}(x_j)$, and $\dot{p}(x_j,W(x_j))=\bar{p}(x_j)$. These are the values of $\dot{u}_i$, $\dot{\rho}$, $\dot{s}$, and $\dot{p}$ on a wave front, and there are two such functions which we define in this way, one corresponding to each side of the wave front. The derivatives of either of these functions are given by

$$\begin{aligned}(\bar{u}_i)_{x_j}&=(\dot{u}_i)_{x_j}+\dot{u}_{i_t}W_{x_j},\\ \bar{\rho}_{x_j}&=\dot{\rho}_{x_j}+\dot{\rho}_tW_{x_j},\\ \bar{s}_{x_j}&=\dot{s}_{x_j}+\dot{s}_tW_{x_j}.\end{aligned} \tag{43}$$

Combining (39)–(42) with (43) yields

$$\dot{\rho}_t[1-u_iW_{x_i}]+\dot{u}_{i_t}[-\rho W_{x_i}]=-\rho(\bar{u}_i)_{x_i}-\rho_{x_i}\bar{u}_i-\bar{p}(u_i)_{x_i}-\bar{p}_{x_i}u_i, \tag{44}$$

$$\dot{\rho}_t[-p_\rho\rho^{-1}W_{x_i}]+\dot{u}_{i_t}[1-u_jW_{x_j}]+\dot{s}_t[-p_s\rho^{-1}W_{x_i}]=\dot{f}_i-p_\rho\rho^{-1}\bar{\rho}_{x_i}-p_s\rho^{-1}\bar{s}_{x_i}+\bar{\rho}\rho^{-2}p_{x_i}-(u_i)_{x_j}\bar{u}_j-u_j(\bar{u}_i)_{x_j}-\rho^{-1}(p_\rho)_{x_i}\bar{\rho}-\rho^{-1}(p_\rho)_{x_i}\bar{s}, \tag{45}$$

$$\dot{s}_t[1-u_jW_{x_j}]=-\bar{u}_js_{x_j}-u_j\bar{s}_{x_j}. \tag{46}$$

One such set of equations is obtained on each side of the wave front. We now subtract each equation on side two from the corresponding equation on side one. Then, remembering that the basic flow was assumed to be continuous, and that $\Delta\dot{s}=\Delta\bar{s}=0$, we obtain

$$[1-u_jW_{x_j}]\Delta(\dot{\rho}_t)+[-\rho W_{x_i}]\Delta(\dot{u}_{i_t})$$
$$=-\rho\Delta(\bar{u}_i)_{x_i}-\rho_{x_i}\Delta\bar{u}_i-(u_i)_{x_i}\Delta\bar{\rho}-u_i\Delta\bar{\rho}_{x_i}, \quad (47)$$

$$[-p_\rho\rho^{-1}W_{x_i}]\Delta(\dot{\rho}_t)$$
$$+[1-u_jW_{x_j}]\Delta(\dot{u}_{i_t})+[-p_s\rho^{-1}W_{x_i}]\Delta(\dot{s}_t)$$
$$=[\rho^{-2}p_{x_i}-\rho^{-1}(p_\rho)_{x_i}]\Delta\bar{\rho}-p_\rho\rho^{-1}\Delta\bar{\rho}_{x_i}$$
$$-(u_i)_{x_j}\Delta\bar{u}_j-u_j\Delta(\bar{u}_i)_{x_j}, \quad (48)$$

$$[1-u_jW_{x_j}]\Delta(\dot{s}_t)=-s_{x_j}\Delta\bar{u}_j. \quad (49)$$

Equations (47)–(49) are a set of five linear inhomogeneous equations for the five quantities $\Delta\dot{\rho}_t$, $\Delta\dot{u}_{i_t}$, and $\Delta\dot{s}_t$. Let us compare these equations with the shock conditions which are a set of five homogeneous equations for $\Delta\dot{\rho}$, $\Delta\dot{u}_i$, and $\Delta\dot{s}$. These latter equations can be rewritten in the form

$$U\Delta\dot{\rho}+\rho\nu_i\Delta\dot{u}_i=0, \quad (50)$$

$$\nu_ip_\rho\Delta\dot{\rho}+\rho U\Delta\dot{u}_i+\nu_ip_s\Delta\dot{s}=0, \quad (51)$$

$$\rho U\epsilon_s\Delta\dot{s}=0. \quad (52)$$

If (47) is multiplied by $1/(W_k^2)^{\frac{1}{2}}$, (48) by $\rho/(W_k^2)^{\frac{1}{2}}$, and (49) by $\rho\epsilon_s/(W_k^2)^{\frac{1}{2}}$ then the coefficient matrix on the left side becomes identical with that of (50)–(52). Therefore, (47)–(49) can have a solution if and only if the right sides of these equations satisfy a certain orthogonality condition, since by assumption the homogeneous equations have a nonzero solution. To obtain this condition we eliminate $\Delta\dot{s}_t$ from (48) by means of (49), and similarly eliminate $\Delta\dot{s}$ from (51) by means of (52). The resulting sets of four equations in four unknowns will both have the same coefficient matrix, which becomes symmetric if we multiply (47) and (50) by p_ρ and (48), (51) by ρ. Now the solvability condition for (47), (50) is that the right side be orthogonal to the solution of the transposed homogeneous equations, which are just (50), (51) because of the symmetry. Thus we obtain the condition

$$-p_\rho\{\rho(\Delta\bar{u}_i)_{x_i}+\rho_{x_i}\Delta\bar{u}_i+(u_i)_{x_i}\Delta\bar{\rho}+u_i\Delta\bar{\rho}_{x_i}\}\Delta\dot{\rho}$$
$$+\rho^2\{[\rho^{-2}p_{x_i}-\rho^{-1}(p_\rho)_{x_i}]\Delta\bar{\rho}-p_\rho\rho^{-1}\Delta\bar{\rho}_{x_i}$$
$$-(u_i)_{x_j}\Delta\bar{u}_j-u_j(\Delta\bar{u}_i)_{x_j}\}\Delta\dot{u}_i$$
$$+U^{-1}\rho p_s s_{x_j}\Delta\bar{u}_j\nu_i\Delta\dot{u}_i=0. \quad (53)$$

Equation (53) can be considerably simplified and solved explicitly. First, we note that $\Delta\bar{u}_i=\Delta\dot{u}_i$ and $\Delta\bar{\rho}=\Delta\dot{\rho}$. Then we eliminate $\Delta\dot{u}_i$ in terms of $\Delta\dot{p}$ by (24), and we eliminate $\Delta\dot{\rho}$ by the relation $\Delta\dot{\rho}=c^2\Delta\dot{p}$, which follows from (5) and (28). In this way we obtain from (53) an equation in which the only acoustic quantity appearing is $\Delta\dot{p}$. This equation is, after dividing by $\Delta\dot{p}$,

$$\nu_i\left(\frac{\Delta\dot{p}}{U}\right)_{x_i}+(\nu_i)_{x_i}\frac{\Delta\dot{p}}{U}-(u_i)_{x_i}\frac{\Delta\dot{p}}{c^2}-\frac{u_i(\Delta\dot{p})_{x_i}}{c^2}+\frac{2u_i\Delta\dot{p}c_{x_i}}{c^3}$$
$$-\frac{p_{x_i}\nu_i\Delta\dot{\rho}}{\rho U^3}+\frac{\nu_i(\Delta\dot{p})_{x_i}}{U}-\frac{(u_i)_{x_j}\nu_i\nu_j\Delta\dot{p}}{U^2}$$
$$-\nu_iu_j(\nu_i)_{x_j}\frac{\Delta\dot{p}}{U^2}-\frac{\rho u_j}{U}\left(\frac{\Delta\dot{p}}{\rho U}\right)_{x_j}+p_ss_{x_j}\nu_j\frac{\Delta\dot{p}}{\rho U^3}=0. \quad (54)$$

Although (54) appears to be a partial differential equation for $\Delta\dot{p}$, it can be rewritten as an ordinary differential equation along a ray. To accomplish this we choose

$$\lambda=-1/2c^2(p_j^2)^{\frac{1}{2}}$$

in (38) so that the equation of a ray becomes

$$\frac{dx_i}{d\sigma}=\nu_i-\frac{u_i}{c}. \quad (55)$$

Then for any function g we have $g_{x_i}[\nu_i-(u_i/c)]=dg/d\sigma$, where $dg/d\sigma$ denotes differentiation with respect to σ along a ray. Making use of this fact, and observing that the ninth term in (54) vanishes since $\nu_i^2=1$ and thus $\nu_i(\nu_i)_{x_j}=0$, (54) becomes

$$\frac{d}{d\sigma}\log\frac{(\Delta\dot{p})^2}{\rho c}+\left(\nu_i-\frac{u_i}{c}\right)_{x_i}$$
$$-u_i\left(\frac{1}{c}\right)_{x_i}-\frac{\nu_i\nu_j}{c}(u_i)_{x_j}=0. \quad (56)$$

In obtaining (56) from (54) it was necessary to divide by $\Delta\dot{p}/U$ and to make use of (29) in the form $U=c$.

Equation (56) can be solved at once, with the result

$$\frac{(\Delta\dot{p})^2}{\rho c}=\left[\frac{(\Delta\dot{p})^2}{\rho c}\right]_{\sigma_0}\exp\int_{\sigma_0}^{\sigma}\left[-\left(\nu_i-\frac{u_i}{c}\right)_{x_i}\right.$$
$$\left.+u_i\left(\frac{1}{c}\right)_{x_i}+\frac{\nu_i\nu_j}{c}(u_i)_{x_j}\right]d\sigma. \quad (57)$$

The first exponential factor

$$\exp\int_{\sigma_0}^{\sigma}[-(\nu_i)_{x_i}]d\sigma$$

has a simple geometrical significance when $u_i=0$. In this case σ denotes arc length along a ray, as is shown by (55). Then by following the method of Luneburg[3]

[3] See reference 2, p. 86.

we find that

$$\exp\int_{\sigma_0}^{\sigma}(-v_i)_{x_i}d\sigma=\frac{1}{K},$$

where K is the expansion ratio along the ray. This expansion ratio is the limit of the ratio of the normal cross-sectional area of a tube of rays at σ divided by the corresponding area at σ_0, as the area tends to zero. Of course, the tube of rays is supposed to contain the ray in question, and to converge to it. Thus, when $u_i=0$, (57) becomes

$$\frac{(\Delta\dot{p})^2}{\rho c}=\frac{1}{K}\left[\frac{(\Delta\dot{p})^2}{\rho c}\right]_{\sigma_0}. \tag{58}$$

Since $\Delta\dot{p}/\rho c$ is the normal jump in velocity across an acoustic shock and $\Delta\dot{p}\Delta\dot{u}$ represents the power per unit area in the shock, (58) asserts the conservation of shock front energy in a tube of rays.

Equation (58) yields the shock strength $\Delta\dot{p}$ at any point on a ray in terms of its value at one point on the ray. Once $\Delta\dot{p}$ is known, $\Delta\dot{\rho}$ and $\Delta\dot{u}_i$ can be immediately computed. Since (58) just inyolves K, $\Delta\dot{p}$ can be found once the rays are determined. It is useful to note that the normal cross section of a tube of rays is a portion of a wave front, and that the area of such a portion is proportional to the Gaussian curvature of the surface. Thus K is also the ratio of the Gaussian curvature at a point on a wave front, to the corresponding curvature of another wave front of the same family at a point on the same ray.

E. Reflection and Transmission of an Acoustic Shock at a Contact Discontinuity

We will now consider the reflection and transmission of an acoustic shock at a contact discontinuity in the basic flow. An incident wave front will, in general, give rise to one reflected and one transmitted wave front, and to a modified motion of the contact discontinuity. If these moving surfaces are represented in four dimensional x_i, t space, they become stationary, three-dimensional hypersurfaces. They are joined together on the contact discontinuity surface Γ, which we will represent parametrically, assuming Γ to be stationary, by the equations

$$x_i=g_i(\xi,\eta). \tag{59}$$

The three wave fronts are given by

$$W^I-t=0,\quad W^R-t=0,\quad W^T-t=0. \tag{60}$$

Here, and below, I, R, T will denote incident, reflected, and transmitted, respectively. In order that these surfaces be joined together on Γ it is necessary that $W^I=W^R=W^T$ on Γ, that is,

$$W^I(g_i)=W^R(g_i)=W^T(g_i). \tag{61}$$

This is an identity in ξ, η, and may, therefore, be differentiated with respect to ξ, η, yielding

$$\begin{aligned}&(W_{x_i}{}^I-W_{x_i}{}^R)g_{i\xi}=0,\quad (W_{x_i}{}^I-W_{x_i}{}^R)g_{i\eta}=0,\\&(W_{x_i}{}^I-W_{x_i}{}^T)g_{i\xi}=0,\quad (W_{x_i}{}^I-W_{x_i}{}^T)g_{i\eta}=0.\end{aligned} \tag{62}$$

These equations show that the difference between the normals to W^I and W^R, and the difference between the normals to W^I and W^T are parallel to the normal to Γ.

We will rewrite these conditions after introducing the wave normals p^I, p^R, and p^T, and the unit normal N to Γ. Then the preceding equations imply

$$\begin{aligned}p^R&=p^I+a_1N,\\p^T&=p^I+a_2N.\end{aligned} \tag{63}$$

Here a_1, a_2 are scalars. From (63) we see that p^I, p^R, p^T, and N all lie in one plane, called the plane of incidence. Forming the cross product of (63) with N, we obtain

$$p^R\times N=p^I\times N=p^T\times N. \tag{64}$$

If we now consider magnitudes in the above equations, and recall that the length of a wave normal is the reciprocal of the wave-front velocity along the normal, we obtain

$$\frac{1}{v^R}\sin\theta^R=\frac{1}{v^I}\sin\theta^I=\frac{1}{v^T}\sin\theta^T. \tag{65}$$

The first equality is the law of reflection, and yields $\theta^R=\pi-\theta^I$ if $v^R=v^I$, which is the case when $u_i=0$. The second equality is Snell's law of refraction. The angles of incidence θ^I, reflection θ^R, and transmission θ^T are between p^I, p^R, p^T, and N. Since

$$\theta^T=\sin^{-1}[(v^T/v^I)\sin\theta^I],$$

there is no value of θ^T, and, thus, no transmitted wave front, if $(v^T/v^I)\sin\theta^I>1$, which is called the case of total reflection.

The wave fronts and Γ divide space into five regions in the neighborhood of the point set J on which these surfaces are joined together. We assume that as J is approached in each of these regions, all of the acoustic

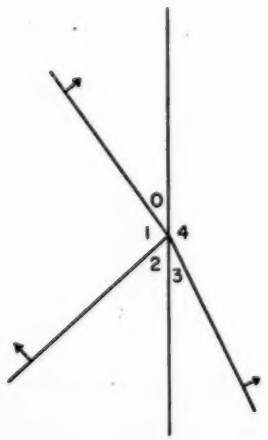


FIG. 1. A shock impinging upon the interface between two media. Reflected and transmitted shocks are shown. The numbers label the various regions between pairs of surfaces and the arrows indicate the directions of motion of the shocks.

quantities have finite limits independent of the direction of approach. If we designate these regions as 0, 1, 2, 3, 4 (see Fig. 1), then the limit of each acoustic quantity will be indicated by a superscript corresponding to the region in which the limit is approached. We next introduce the jumps

$$\Delta u_i^0 = u_i^4 - u_i^0, \quad \Delta u_i^1 = u_i^0 - u_i^1, \quad \Delta u_i^2 = u_i^1 - u_i^2,$$
$$\Delta u_i^3 = u_i^2 - u_i^3, \quad \Delta u_i^4 = u_i^3 - u_i^4, \text{ etc.} \tag{66}$$

It is assumed that Δu_i^0, Δp^0, $\Delta \rho^0$, Δs^0 are known, since these are the jumps across the incident wave front. By applying the jump conditions, we wish to determine the other jumps, especially the reflected and transmitted ones Δu_i^1, Δu_i^3 in terms of these.

From the shock condition (24) at an acoustic shock, we have

$$\Delta u_i^0 = -\nu_i^I (\rho c)^{-1} \Delta p^0, \quad \Delta u_i^1 = -\nu_i^R (\rho c)^{-1} \Delta p^1,$$
$$\Delta u_i^3 = -\nu_i^T (\rho' c')^{-1} \Delta p^3. \tag{67}$$

Here ρ, c designate quantities on the "incident" side, and ρ', c' on the "transmitted" side of the contact discontinuity. Also, from the definitions of the jumps, we have

$$\Delta p^0 + \Delta p^1 + \Delta p^2 + \Delta p^3 + \Delta p^4 = 0, \tag{68}$$

$$\Delta u_i^0 + \Delta u_i^1 + \Delta u_i^2 + \Delta u_i^3 + \Delta u_i^4 = 0. \tag{69}$$

Since pressure is continuous across a contact discontinuity, $\Delta p^2 = \Delta p^4 = 0$, and (68) becomes

$$\Delta p^0 + \Delta p^1 + \Delta p^3 = 0. \tag{70}$$

Now, we take the component of (69) normal to Γ and recall that the normal velocity is continuous across a contact discontinuity. Then we have $N_i \Delta u_i^2 = N_i \Delta u_i^4 = 0$ and (69) becomes, after (67) is used,

$$(\rho c)^{-1} \cos\theta^I \Delta p^0 + (\rho c)^{-1} \cos\theta^R \Delta p^1 + (\rho' c')^{-1} \cos\theta^T \Delta p^3 = 0. \tag{71}$$

Equations (70) and (71) are two equations for the reflected pressure jump Δp^1 and the transmitted pressure jump $-\Delta p^3$. The solution of these equations is

$$\Delta p^1 = \Delta p^0 \begin{vmatrix} \dfrac{\cos\theta^I}{\rho c} & \dfrac{\cos\theta^T}{\rho' c'} \\ \dfrac{\cos\theta^T}{\rho' c'} & \dfrac{\cos\theta^R}{\rho c} \end{vmatrix}, \tag{72}$$

$$-\Delta p^3 = \Delta p^0 \begin{vmatrix} \dfrac{\cos\theta^I}{\rho c} & \dfrac{\cos\theta^R}{\rho c} \\ \dfrac{\cos\theta^T}{\rho' c'} & \dfrac{\cos\theta^R}{\rho c} \end{vmatrix}. \tag{73}$$

The coefficients of Δp^0 in (72) and (73) are the reflection and transmission coefficients, respectively.

If $u_i = 0$ on the incident side of the contact discontinuity then $v^R = v^I$ and (65) shows that $\theta^R = \pi - \theta^I$, provided N is oriented so that $0 \leq \theta^I \leq \pi/2$. Then $\cos\theta^R = -\cos\theta^I$, and (72), (73) become

$$\Delta p^1 = \frac{1-Z}{1+Z} \Delta p^0, \tag{74}$$

$$-\Delta p^3 = \frac{2}{1+Z} \Delta p^0; \quad Z = \frac{\rho c \cos\theta^T}{\rho' c' \cos\theta^I}. \tag{75}$$

Equations (74) and (75) are Fresnel's formulas, usually deduced for reflection of periodic plane waves from a plane interface.

F. Expansion Ratio for Straight Rays

The expansion ratio K was introduced in (58) and determines the variation of the shock strength along a ray. It may be defined as the ratio of the Gaussian curvature of a wave front at one point on a ray to that at another point on the same ray. In a uniform basic flow with constant sound speed, the rays are straight lines orthogonal to the wave fronts. In this case the various wave fronts of a family of wave fronts are parallel surfaces, according to the definition of parallel surfaces employed in differential geometry. If R_1, R_2 are the principal radii of curvature of a surface at a point, then R_1+Z, R_2+Z are the corresponding radii of a parallel surface at a point at distance Z away on the same orthogonal (i.e., ray). Thus,

$$K^{-1} = \frac{R_1 R_2}{(R_1+Z)(R_2+Z)}. \tag{76}$$

From (76) we see that K^{-1} becomes infinite at two points $Z = -R_1$, $-R_2$, called conjugate points, on each ray. The locus of these points for a given family of wave fronts is called a caustic surface, which usually has two branches. The branches touch if on some ray $R_1 = R_2$. A point at which the two branches of the caustic surface touch is usually called a focus. Since K^{-1} becomes infinite on a caustic, the acoustic shock strength also becomes infinite there.

The expansion ratio has also been computed for a uniform basic flow with a curved contact discontinuity separating two regions of different constant sound speed. In this case the expansion ratio for reflected or refracted (i.e., transmitted) wave fronts becomes important. For an incident spherical wave front these expansion ratios were computed by J. B. and H. B. Keller.[4] For an arbitrary incident wave front the result is much more complicated, and is given by Stanley

[4] J. B. Keller and H. B. Keller, J. Opt. Soc. Am. **40**, 48 (1950).

Preiser and Joseph B. Keller.[5] Some parts of these results have also been given by others.

G. Example: The Shock Tube

As an example of the utility of the preceding theory, we will apply it to the shock-tube or one-dimensional explosion. The results are also useful in understanding some of the phenomena in spherical explosions.

We suppose that all quantities depend upon x and t only and that only the x component of velocity, which we call u, differs from zero. The region $-a<x<0$ is supposed to be filled with a perfect, nonconducting gas at high pressure, and bounded at $x=-a$ by a rigid wall. The region $x>0$ is supposed to be filled by a perfect, nonconducting gas at lower pressure. As the basic flow we take the state of rest: $u=0$, $p=\text{const}$, $\rho=\rho_1$ or ρ_2, $s=s_1$ or s_2, according as $x>0$ or $x<0$ (see Fig. 2).

We now assume that at $t=0$ the acoustic quantities vanish for $x>0$ and that for $-a<x<0$ they have the constant values $\dot{p}$, $\dot{\rho}$, $\dot{s}$, $\dot{u}=0$. Thus, at $t=0$ the surface $x=0$ is an acoustic discontinuity. By Huygen's construction we see that, if this gives rise to acoustic shocks at later times they will be located at $x=c_1t$ or $-c_2t$ where c_1, c_2 are the sound speeds in the basic flow for $x>0$ and $x<0$, respectively. Furthermore, the strengths of these shocks will be constant since the expansion ratio for a plane wave front is unity.

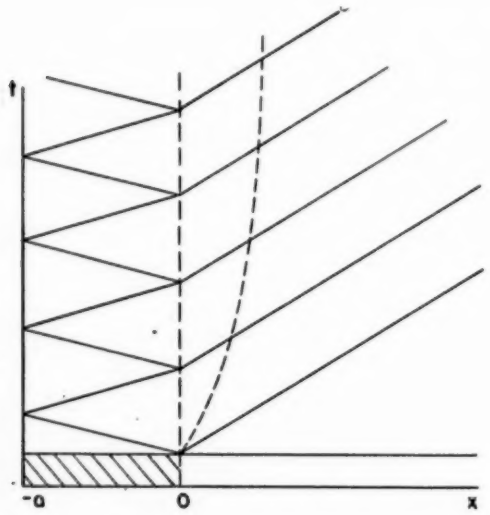


FIG. 2. An $x-t$ diagram of the shock tube. The x axis is horizontal, and at $t=0$ the high-pressure gas lies between a rigid wall at $x=-a$ and a diaphragm at $x=0$, while the low-pressure gas lies to the right of $x=0$. At $t=0$ the diaphragm is broken and a shock travels to the right while a rarefaction travels to the left (solid lines). The rarefaction is reflected from the wall at $x=-a$, travels to the right, and is reflected and transmitted at $x=0$, which is the undisturbed boundary between the two gases. This process repeats periodically. The dashed line on the right is the location of the moving interface between the two gases. This interface moves with constant velocity between successive shock reflections. The vertical dashed line is the undisturbed location of the interface.

To solve the problem, we assume that the solution is piecewise constant, the discontinuities occurring across the shocks and across the contact discontinuity at $x=0$. The acoustic quantities to the right of the shock $x=c_1t$ and to the left of the shock $x=-c_2t$ have their initial values, while between the shocks the acoustic pressure and velocity will be denoted by $\bar{p}$, $\bar{u}$. These quantities are continuous across the contact discontinuity $x=0$, and we assume that they are constant.

From the shock conditions applied at each shock we find

$$\frac{\bar{p}-\dot{p}}{\rho_2 c_2}=\bar{u}=\frac{\bar{p}}{\rho_1 c_1}. \tag{77}$$

Thus,

$$\bar{p}=\dot{p}\frac{(\rho c)_1}{(\rho c)_1+(\rho c)_2}, \tag{78}$$

$$\bar{u}=\dot{p}\frac{1}{(\rho c)_1+(\rho c)_2}. \tag{79}$$

The foregoing results (78), (79) will apply until the left shock reaches the rigid wall at $x=-a$, when a reflected shock is introduced. Since $\bar{u}=0$ at a rigid wall, the velocity jump across the reflected shock is the negative of that behind the incident one. Therefore, the reflected pressure jump is equal to the incident pressure jump. This result also follows from the reflection formula (74) when $\rho'c'$ is infinite. The solution with one reflected shock applies until the reflected shock reaches the contact discontinuity, when it is partially reflected and partially transmitted according to (74), (75). The reflected shock will again be reflected at $x=-a$, and this process will continue indefinitely.

The strength of the first shock traveling to the right of $x=0$ is given by (78), since the acoustic pressure in front of it is zero. We will denote the strength of this shock by $\alpha\dot{p}$, where $\alpha=(\rho c)_1[(\rho c)_1+(\rho c)_2]^{-1}$. Then the strength of the first shock traveling to the left is $(\alpha-1)\dot{p}$, which is also the strength of this shock after reflection from $x=-a$. The shock transmitted through the contact discontinuity at $x=0$ will have strength $T(\alpha-1)\dot{p}$, and the reflected shock will have strength $R(\alpha-1)\dot{p}$. Here R and T are the reflection and transmission coefficients given by (74), (75) with $\theta^I=0$, $\theta^R=\pi$. In this way we see that the strength of the nth shock traveling to the right of $x=0$ is $R^{n-2}T(\alpha-1)\dot{p}$ for $n>1$.

The pressure p_n behind the nth shock traveling to the right of $x=0$ can easily be computed by summing all the pressure jumps across the preceding shocks. In

[5] S. Preiser and J. B. Keller, Research Report No. EM-20, Washington Square College Math. Research Group, New York University, New York (1950).

this way we find $p_1=\alpha\dot{p}$ and

$$p_n=\left[\alpha+T(\alpha-1)\frac{1-R^{n-1}}{1-R}\right]\dot{p}=2R^{n-1}\dot{p} \quad n\geq 2. \quad (80)$$

The velocity u_n behind the nth shock traveling to the right of $x=0$ is then given by

$$u_n=p_n/\rho_1 c_1. \quad (81)$$

The contact discontinuity, at $x=0$ in the basic flow, moves a distance $d_n=(2a/c_2)u_n$ between the nth and $(n+1)$st shocks. Thus, the total distance D_n which it moves until the arrival of the $(n+1)$st shock is

$$D_n=\frac{2a}{c_2}\cdot\frac{\dot{p}}{\rho_1 c_1}\left[n\alpha+\frac{T(\alpha-1)}{1-R}(n-1-\sum_{k=1}^{n-1}R^k)\right]$$

$$=\frac{a\dot{p}}{\rho_2 c_2^2}(1-R^n). \quad (82)$$

From (82) we see that D_n oscillates if $R<0$, and D_n varies monotonically with n if $R>0$.

JOURNAL OF APPLIED PHYSICS VOLUME 25, NUMBER 8 AUGUST, 1954

Transient Wave Analysis in a Linear Time-Dependent Material*

R. D. Glauz† and E. H. Lee‡
Brown University, Providence, Rhode Island

(Received August 20, 1953)

We consider a body which under stress exhibits an elastic, a viscous, and a retarded elastic response. For the particular case of a constant velocity applied at the end of a semi-infinite bar, the transient stresses, strains and velocities are found by the method of characteristics. From this solution can also be obtained the stresses for the case of a constant stress applied at the end of a semi-infinite bar. By superposition we can also solve problems involving finite bars. For a constant velocity applied at the end of the bar an integral expression is derived for the stress at that end as a function of time.

1. FUNDAMENTAL CONSIDERATIONS

LEE and Kanter[1] found transient stresses and strains for a body whose stress-strain law could be represented by a two parameter Maxwell model. We consider here a stress-strain law represented by a four parameter model as in Fig. 1. The stress-strain law is given by

$$\begin{aligned}\epsilon&=\epsilon_1+\epsilon_2+\epsilon_3\\ E_1\epsilon_1&=\sigma\\ E_2\epsilon_2+\frac{1}{\mu_2}\epsilon_{2,t}&=\sigma\\ \frac{1}{\mu_3}\epsilon_{3,t}&=\sigma,\end{aligned} \quad (1)$$

where the subscript t denotes differentiation partially with respect to the time t. The correspondence to the model arises by considering σ as the force applied at the point A causing an extension of the model which will be the strain ϵ. ϵ_1 will correspond to the extension of the spring, ϵ_2 to the extension of the combination spring and dashpot in parallel, and ϵ_3 to the extension of the dashpot.[2]

With the four parameter model we have added a retarded elastic element to the two parameter Maxwell model. With this added response it will be, in general, possible to approximate more closely the actual behavior of a polymer.

When a stress is applied to an amorphous high polymer there are microscopic and macroscopic motions. There is an instantaneous elastic response wherein the atoms move further apart, valence angles are changed, or valence bonds are stretched or compressed. There is flow as the molecules shift their centers of gravity and reach new equilibrium positions. There is also a retarded elastic phenomena. This is the unfolding of the long molecules so that they are oriented in the direction of the applied stress.

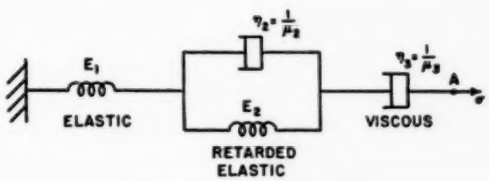


FIG. 1. Four parameter model.

* Sponsored under a Department of Defense Contract, NOrd-11496 between the Bureau of Ordnance, Navy Department, and Brown University.

† Staff member at University of California, Los Alamos Scientific Laboratory. Formerly research associate in the Graduate Division of Applied Mathematics, Brown University.

‡ Professor of Applied Mathematics, Brown University, Providence, Rhode Island.

[1] E. H. Lee and I. Kanter, J. Appl. Phys. **24**, 1115–1122 (1953).

[2] T. Alfred, Jr., *Mechanical Behavior of High Polymers* (Interscience Publishers, Inc., New York, 1948).

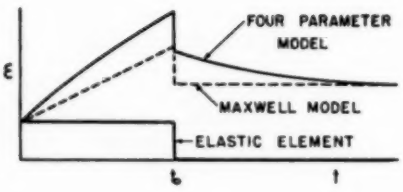


FIG. 2.

When the stress is removed there will be two types of recovery: the instantaneous recovery as the valence angles and valence bonds return to their unstressed equilibrium positions, and the retarded elastic recovery as the molecules which were unfolded and uncurled return to their equilibrium state. The viscous flow is not recoverable as it is a result of the molecules moving to new equilibrium positions.

In Fig. 2 we have compared the four parameter model with the Maxwell model and the elastic element.

Let us first consider a creep test. A load is applied at time $t=0$ and this load maintained until time $t=t_0$. The elastic element is strained instantaneously and maintains the strain as long as the load remains constant. The Maxwell model has the same instantaneous elastic strain but also exhibits a constant flow rate. The four parameter model has the same instantaneous elastic strain and viscous flow but also has a delayed elastic response. The rounding off of the initial portion of the flow is due to this delayed elastic portion of the model which approaches the strain σ/E_2 asymptotically.

Let us now consider unloading at time $t=t_0$. The elastic element immediately returns to zero strain. The Maxwell model recovers the elastic portion of the strain but still has a permanent strain because of the viscous flow. The four parameter model also immediately recovers the elastic portion of the strain and asymptotically recovers the strain of the delayed elastic portion of the model. A permanent deformation is left because of the viscous flow.

Thus we can approximate more closely the actual behavior of an amorphous high polymer with a four parameter model since we can include the motion due to the unfolding and preferential orientation of the molecules out of their equilibrium state.

With a random distribution of molecules there would be an entire spectrum of relaxation times for the reversible elastic portion of the model. If so, we must consider the four parameter model as only an approximation wherein we have picked some average value for a relaxation time which will give satisfactory results for a considered time interval.[2]

In addition to the stress-strain law (1) we have the dynamic equilibrium condition

$$\sigma_x=\rho u_{tt}. \tag{2}$$

We also use the definitions of ϵ, v

$$\epsilon=\frac{\partial u}{\partial x}, \quad v=\frac{\partial u}{\partial t}. \tag{3}$$

Equations (1), (2), and (3) comprise the system of equations to be solved in conjunction with appropriate boundary conditions. If we put these equations in dimensionless form we obtain an equivalent problem to solve in which the five physical constants have been combined to yield two parameters. Eliminatng $i\epsilon_1$, ϵ_3 from (1), (2), and (3) we put the resulting equations in the dimensionless form

$$\left.\begin{aligned} \epsilon_2^*+\epsilon_{2,\tau}^*&=\sigma^*,\\ \epsilon_\tau^*&=v_\xi^*,\\ \sigma_\xi^*&=v_\tau^*,\\ \epsilon_\tau^*-\epsilon_{2,\tau}^*-\alpha\sigma^*&=\beta\sigma_\tau^*, \end{aligned}\right\} \tag{4}$$

where

$$\left.\begin{aligned} x&=\frac{1}{\mu_2(\rho E_2)^{\frac{1}{2}}}\xi, \quad t=\frac{1}{\mu_2 E_2}\tau,\\ \epsilon&=\frac{E_1}{E_2}\epsilon^*, \quad \epsilon_2=\frac{E_1}{E_2}\epsilon_2^*, \quad \sigma=E_1\sigma^*, \quad v=\frac{E_1}{(\rho E_2)^{\frac{1}{2}}}v^*,\\ \alpha&=\frac{\mu_3}{\mu_2}, \quad \beta=\frac{E_2}{E_1}. \end{aligned}\right\} \tag{5}$$

By simple elimination of variables we can find the partial differential equations satisfied by the dimensionless stress, strain, and velocity.[3] It is found that these three satisfy the same partial differential equation so that if we find the transient stresses for a constant stress applied at the end of a bar we will also have the solution of the problem of finding the transient velocities for a constant velocity applied at the end of the bar.

We are primarily interested in the transient stresses which occur in the body and we can find the most information about the stresses if we find the stress and velocity distribution for a constant velocity applied at the end of the bar. As mentioned above, we have thus automatically solved the problem of finding the stresses for a constant stress applied. We note that in the characteristic solution the values of stress and velocity at a point are found simultaneously in constructing the solution.

The numerical characteristic solution is found for values of $\alpha=1$, $\beta=1$ in which case all elements of the four parameter model have a significant effect. A check of the characteristic solution has been found in the form of an integral which represents the stress at the end of the bar where the constant velocity is applied.

[3] See Appendix I for the partial differential equations.

2. CHARACTERISTIC EQUATIONS

The system of Eqs. (1) is a hyperbolic system of linear partial differential equations and can be numerically integrated by the method of characteristics under appropriate boundary conditions. We can immediately eliminate ϵ_τ^* from (4) to give

$$\begin{aligned} &\epsilon_{2,\tau}{}^* = \sigma^* - \epsilon_2{}^* \\ &\sigma_\xi{}^* - v_\tau{}^* = 0 \\ &\epsilon_{2,\tau}{}^* + \beta\sigma_\tau{}^* - v_\xi{}^* = -\alpha\sigma^*. \end{aligned} \tag{6}$$

The dependent variables σ^*, ϵ_2^*, and v^* are functions of both ξ and τ and can be considered graphically as surfaces in space, e.g., σ^*, ξ, τ space. The functions σ^*, ϵ_2^*, v^* are well behaved in the interior of the region bounded by the ξ axis and the line $(\beta\xi)^{\frac{1}{2}} - \tau = 0$, where β has been defined in (5). Consequently their derivatives will take on particular values as we move along some curve in this region.

We have

$$\begin{aligned} d\sigma^* &= \frac{\partial\sigma^*}{\partial\xi}d\xi + \frac{\partial\sigma^*}{\partial\tau}d\tau \\ d\epsilon_2{}^* &= \frac{\partial\epsilon_2{}^*}{\partial\xi}d\xi + \frac{\partial\epsilon_2{}^*}{\partial\tau}d\tau \\ dv^* &= \frac{\partial v^*}{\partial\xi}d\xi + \frac{\partial v^*}{\partial\tau}d\tau. \end{aligned} \tag{7}$$

The systems of Eqs. (6) and (7) together form a set of six equations in the six unknowns $\partial\sigma^*/\partial\xi$, $\partial\sigma^*/\partial\tau$, $\partial\epsilon_2{}^*/\partial\xi$, $\partial\epsilon_2{}^*/\partial\tau$, $\partial v^*/\partial\xi$, $\partial v^*/\partial\tau$, and we can solve for each of them. For example, using determinants to solve for $\partial\sigma^*/\partial\xi$ we have

$$\frac{\partial\sigma^*}{\partial\xi} = \frac{\begin{vmatrix} \sigma-\epsilon_2{}^* & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \\ -\alpha\sigma^* & \beta & 0 & 1 & -1 & 0 \\ d\sigma^* & d\tau & 0 & 0 & 0 & 0 \\ d\epsilon_2{}^* & 0 & d\xi & d\tau & 0 & 0 \\ dv^* & 0 & 0 & 0 & d\xi & d\tau \end{vmatrix}}{\begin{vmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & -1 \\ 0 & \beta & 0 & 1 & -1 & 0 \\ d\xi & d\tau & 0 & 0 & 0 & 0 \\ 0 & 0 & d\xi & d\tau & 0 & 0 \\ 0 & 0 & 0 & 0 & d\xi & d\tau \end{vmatrix}} \tag{8}$$

The characteristics are defined as those curves in the $\xi-\tau$ plane along which the derivatives $\partial\sigma^*/\partial\xi$, $\partial\sigma^*/\partial\tau$, $\partial\epsilon_2{}^*/\partial\xi$, $\partial\epsilon_2{}^*/\partial\tau$, $\partial v^*/\partial\xi$, $\partial v^*/\partial\tau$ are not uniquely determined.[4] For $\partial\sigma^*/\partial\xi$ to be finite and indeterminate we must have both determinants of (8) equal to zero. Setting the denominator of (8) equal to zero yields the equation determining the characteristics:

$$d\xi[(\beta)^{\frac{1}{2}}d\xi + d\tau][(\beta)^{\frac{1}{2}}d\xi - d\tau] = 0, \tag{9}$$

from which we obtain the equations of the three sets of characteristics:

$$\begin{aligned} &\xi = \text{constant} \\ &\xi + \frac{1}{(\beta)^{\frac{1}{2}}}\tau = \text{constant} \\ &\xi - \frac{1}{(\beta)^{\frac{1}{2}}}\tau = \text{constant}. \end{aligned} \tag{10}$$

Setting the numerator of (8) equal to zero yields the condition

$$d\xi[d\tau dv^* - \beta d\sigma^* d\xi + \epsilon_2{}^* d\xi d\tau - (1+\alpha)\sigma^* d\xi d\tau] = 0. \tag{11}$$

We desire to have this expression identically satisfied on all the characteristics. (11) is obviously satisfied along the set of characteristics $\xi = \text{constant}$. Along the characteristic $\xi - 1/(\beta)^{\frac{1}{2}}\tau = \text{constant}$, we have the necessary condition

$$(\beta)^{\frac{1}{2}}dv^* - \beta d\sigma^* = [(1+\alpha)\sigma^* - \epsilon_2]d\tau. \tag{12}$$

Similarly, on $\xi + 1/(\beta)^{\frac{1}{2}}\tau = \text{constant}$ we have

$$(\beta)^{\frac{1}{2}}dv^* + \beta d\sigma^* = -[(1+\alpha)\sigma^* - \epsilon_2]d\tau. \tag{13}$$

In a similar manner we solve for the other partial derivatives of σ^*, ϵ_2^*, and v^* and obtain the additional information that along the family of characteristics $\xi = \text{constant}$ we must have

$$d\epsilon_2{}^* = (\sigma^* - \epsilon_2{}^*)d\tau. \tag{14}$$

Equation (14) could also have been obtained by integrating the first equation of (4) in the τ direction. In (10) we have the equations of the families of characteristics and using (12), (13), and (14) we can numerically find the values of σ^*, ϵ_2^*, and v^* by integrating along the characteristics from the known data on the boundary. We shall find the boundary conditions which hold along the wave front and also complete the setting up of the problem in section 3.

3. WAVE-FRONT BOUNDARY CONDITIONS

According to the theory of characteristics and the zero initial conditions, a point ξ_0 of the bar will be unstressed until the stress wave reaches that point at time $\tau_0 = (\beta)^{\frac{1}{2}}\xi_0$. At time τ_0 there will be a jump in the stress and velocity at the point ξ_0. White and Griffis[5]

[4] R. Courant and D. Hilbert, *Methoden der Mathematischen Physik* (Verlag von Julius Springer, Berlin, 1937), Vol. 2, pp. 110–122.

[5] M. P. White and L. Griffis, J. Appl. Mech. **69**, A-337–A-343 (1947).

have evaluated this jump at the wave front by equating the impulse and change of momentum. Putting their result in dimensionless form we have

$$\Delta v^* = -(\beta)^{\frac{1}{2}}\Delta\sigma^*. \tag{15}$$

Also, from Malvern[6] we have

$$\Delta v^* = -\frac{1}{(\beta)^{\frac{1}{2}}}\Delta\epsilon^*. \tag{16}$$

We consider the body to be initially unstressed and unstrained, i.e., initially $\sigma^*=\epsilon_2^*=v^*=0$. We can integrate (15) and (16) to obtain the relation holding just after the wave passes

$$\sigma^* = \frac{\epsilon^*}{\beta} = -\frac{v^*}{(\beta)^{\frac{1}{2}}}. \tag{17}$$

From physical considerations, or as implied by the first of (4) for finite σ^*, we see that

$$\epsilon_2^* = 0, \tag{18}$$

just after the wave passes since this is the strain corresponding to the retarded elastic element of the four parameter model. The wave front corresponds to the characteristic $\xi = 1/(\beta)^{\frac{1}{2}}\tau$ in the $\xi-\tau$ plane, and putting (17) and (18) in (12) yields after integration

$$\begin{aligned}
\sigma^* &= c_1 \exp\left[-\left(\frac{1+\alpha}{2\beta}\right)\tau\right] \\
\epsilon^* &= c_1\beta \exp\left[-\left(\frac{1+\alpha}{2\beta}\right)\tau\right] \\
v^* &= -c_1\sqrt{\beta} \exp\left[-\left(\frac{1+\alpha}{2\beta}\right)\tau\right],
\end{aligned} \tag{19}$$

which are the boundary conditions along the wave front.

To construct a characteristic solution we must additionally prescribe either the stress, velocity, or displacement at the end of the bar.

If the stress is prescribed at $\xi=0$ it is easily found that along $\xi=0$

$$\begin{aligned}
\sigma^* &= f_0 \\
\epsilon^* &= f_0(\alpha\tau - e^{-\tau} + 1 + \beta) \\
\epsilon_2^* &= f_0(1-e^{-\tau}) \\
v_\xi^* &= f_0(\alpha + e^{-\tau}).
\end{aligned} \tag{20}$$

Similarly, if the velocity is prescribed at $\xi=0$ we have that on $\xi=0$

$$\begin{aligned}
v^* &= -g_0 \\
\sigma_\xi^* &= 0.
\end{aligned} \tag{21}$$

[6] L. E. Malvern, The Propagation of Longitudinal Waves of Plastic Deformation in a Bar of Material Exhibiting a Strain-rate Effect, Technical Report, A11-39, Graduate Division of Applied Mathematics, Brown University, Providence, Rhode Island (1949).

In the numerical work it is only necessary to use the first equations of (20) or (21) since we have two characteristics with characteristic relations on them along which to evaluate the other two unknown functions. We have computed the characteristic solution for $\alpha=\beta=g_0=1$. For this choice of the parameters all four elements of the model give significant strains. This is in contrast, for example, to a choice of very small α in which the viscous element would be negligible for short or medium time intervals. The choice of unity for g_0 is purely to obtain convenient magnitudes for the dimensionless variables. Since the problem is linear, another choice merely changes the values of all dependent variables proportionately.

From the nature of finding the characteristic solution by working from the origin outward it is desirable to have an accurate analytic representation for σ^*, ϵ_2^*, v^* in the neighborhood of $\xi=\tau=0$. For small ξ, τ a power series in these two variables converges rapidly and gives satisfactory values of the functions. Let us assume that σ^*, ϵ_2^*, v^* have approximate polynomial representations of the form:

$$\begin{aligned}
\sigma^* &= a_1 + a_2\xi + a_3\tau + a_4\xi^2 + a_5\xi\tau \\
&\quad + a_6\tau^2 + a_7\xi^3 + a_8\xi^2\tau + a_9\xi\tau^2 + a_{10}\tau^3 \\
\epsilon_2^* &= b_1 + \cdots + b_{10}\tau^3 \\
v^* &= c_1 + \cdots + c_{10}\tau^3.
\end{aligned} \tag{22}$$

For the case of a constant velocity applied at the end of the bar $\xi=0$ we find by substitution in the partial differential equations, and using the wave-front boundary conditions that

$$\begin{aligned}
\sigma^* &= g_0\left[\frac{1}{(\beta)^{\frac{1}{2}}} - \frac{\gamma}{(\beta)^{\frac{1}{2}}}\tau - \frac{(\gamma^2\beta+1)}{4(\beta)^{\frac{1}{2}}}\xi^2 + \frac{(3\gamma^2\beta+1)}{4\beta(\beta)^{\frac{1}{2}}}\tau^2 - \tfrac{1}{6}\xi^3\right. \\
&\quad \left. + \frac{(\gamma^3\beta+\gamma+1)}{4(\beta)^{\frac{1}{2}}}\xi^2\tau - \frac{(5\gamma^3\beta+3\gamma+1)}{12\beta(\beta)^{\frac{1}{2}}}\tau^3\right] \\
\epsilon_2^* &= g_0\left[-\xi + \frac{1}{(\beta)^{\frac{1}{2}}}\tau - \frac{(\beta)^{\frac{1}{2}}(\gamma-1)}{2}\xi^2 + \xi\tau - \frac{(\gamma+1)}{2(\beta)^{\frac{1}{2}}}\tau^2\right. \\
&\quad + \frac{(2+\alpha-\beta)}{6}\xi^3 - \frac{(\gamma^2\beta+2+\alpha-2\beta)}{4(\beta)^{\frac{1}{2}}}\xi^2\tau - \tfrac{1}{2}\xi\tau^2 \\
&\quad \left. + \frac{(3\gamma^2\beta+2+\alpha+2\beta)}{12\beta(\beta)^{\frac{1}{2}}}\tau^3\right] \\
v^* &= g_0\left[-1 + (\beta)^{\frac{1}{2}}\gamma\xi + \tfrac{1}{2}\xi^2 - \frac{(\gamma^2\beta+1)}{2(\beta)^{\frac{1}{2}}}\xi\tau\right. \\
&\quad \left. + (\beta)^{\frac{1}{2}}\frac{(3-3\gamma-\beta\gamma^3)}{12}\xi^3 - \tfrac{1}{2}\xi^2\tau + \frac{(\gamma^3\beta+\gamma+1)}{4(\beta)^{\frac{1}{2}}}\xi\tau^2\right] \\
\gamma &= \frac{1+\alpha}{2\beta}.
\end{aligned} \tag{23}$$

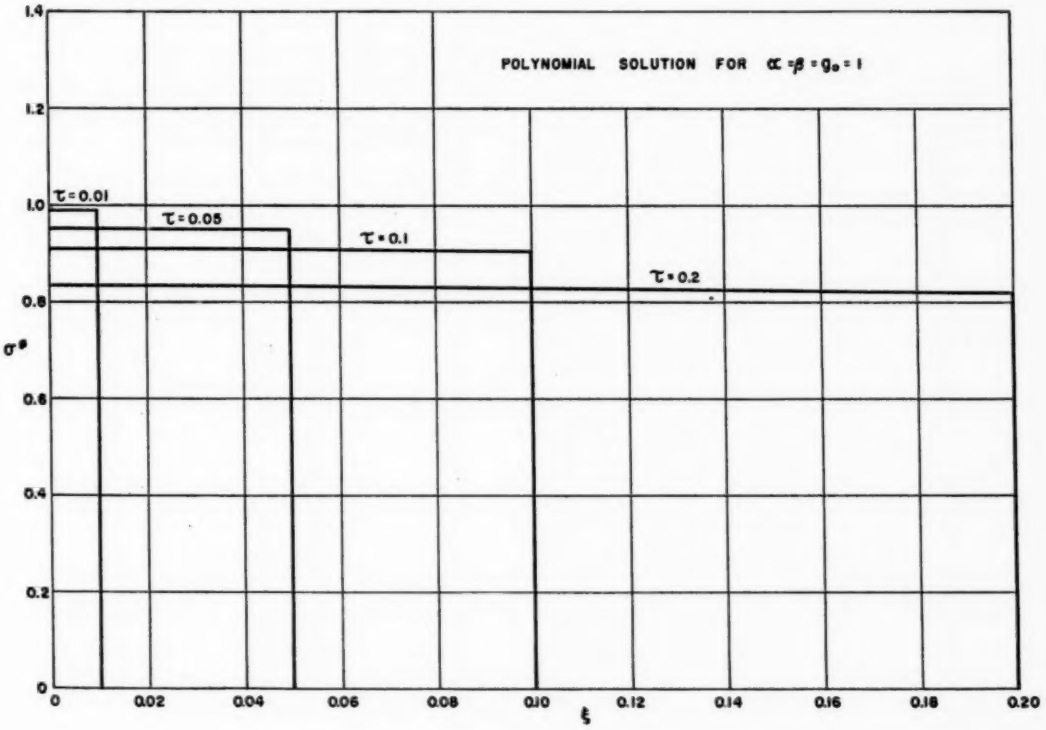


FIG. 3.

In Fig. 3 we have plotted curves found by evaluating σ^* in (23) for the special case of $\alpha=\beta=g_0=1$.

The characteristic solution was built outward using a linear interpolation between points and the following lengths of intervals along $(\beta)^{\frac{1}{2}}\xi-\tau=0$

$$0<\tau\leq 3 \qquad d\tau=0.1$$
$$3<\tau\leq 6 \qquad d\tau=0.2$$
$$6<\tau\leq 10 \qquad d\tau=0.4.$$

The size of the intervals can be increased as the solution proceeds, because the surfaces become smoother and the error of using linear interpolation decreases. Malvern[6] has given a description of the complete numerical method.

The solution is given in the sets of curves in Figs. 4–9. Comparison of these curves with those in reference 1 shows the effect on the stresses and velocities of the addition of the retarded elastic element. For short time the curves are almost identical as the viscous portion of the retarded elastic element has not allowed appreciable motion. For a longer time however the retarded elastic element has flowed relieving some of the stress of the purely elastic element and giving a reversal of curvature to the stress curves.

After a longer time the effect of the retarded elastic element is reduced. This is easily seen from inspection of the graphs of ϵ_2^*. The retarded elastic strain ϵ_2^* at the end of the bar starts at zero and rapidly increases to a maximum. This is the point where the retarded elastic element has its largest smoothing effect on the stress. ϵ_2^* then decreases.

In order to obtain a more complete picture for the four parameter model, wave propagation will have to be evaluated for other values of α and β. However, until measurements of these values for specific materials have been made this seems hardly worthwhile, the present example being presented as a typical case.

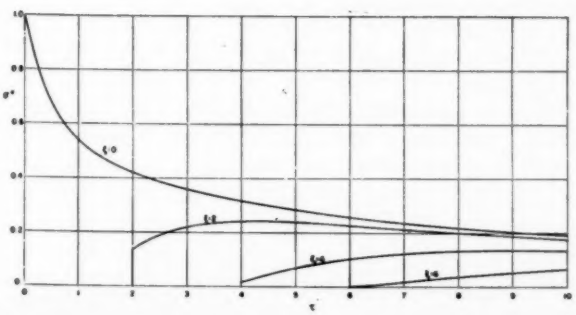


FIG. 4. Characteristic solution for $\alpha=\beta=g_0=1$.

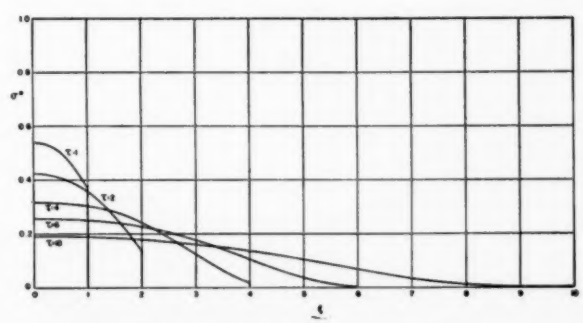


FIG. 5. Characteristic solution for $\alpha=\beta=g_0=1$.

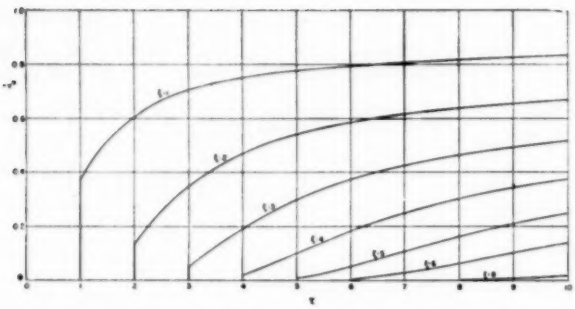


FIG. 6. Characteristic solution for $\alpha=\beta=g_0=1$.

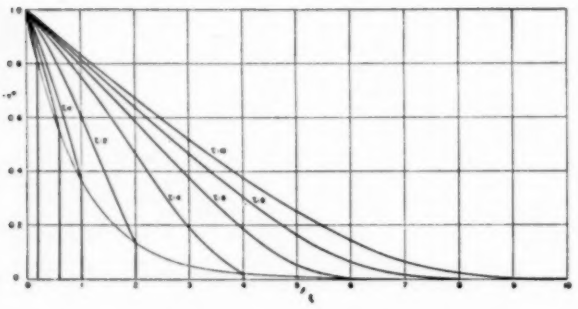


FIG. 7. Characteristic solution for $\alpha=\beta=g_0=1$.

This solution can be utilized to determine the stress distributions in finite rods by means of the image method discussed in reference 1. This development can be handled graphically, and no further computations are necessary.

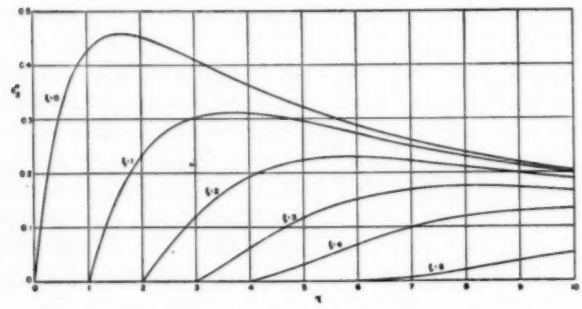


FIG. 8. Characteristic solution for $\alpha=\beta=g_0=1$.

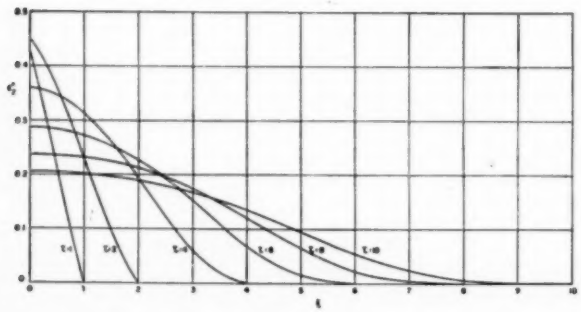


FIG. 9. Characteristic solution for $\alpha=\beta=g_0=1$.

4. ANALYTIC SOLUTION

Several methods of effecting an analytic solution have been tried but the solutions are much too complicated to yield results which can be reasonably evaluated. The characteristic solution has been found to be the quickest method to obtain results. However an analytic expression for the stress at the end of the bar as a function of time has been found to check the characteristic solution.

For the case $\alpha=\beta=1$ we have obtained the following result using the Laplace transform method:

$$\sigma^*(0,\tau)=\exp(-3\tau/2)I_0\left(\frac{(5)^{\frac{1}{2}}}{2}\tau\right)+\frac{1}{2}\int_0^\tau \exp(-\lambda-\tau/2)I_0\left(\frac{(5)^{\frac{1}{2}}}{2}\lambda\right)\times\{I_0[\tfrac{1}{2}(\tau-\lambda)]+I_1[\tfrac{1}{2}(\tau-\lambda)]\}d\lambda. \quad (24)$$

The derivation is given in Appendix II. To check our characteristic solution, (24) has been evaluated for $\tau=10$ and the value 0.194 was obtained as compared to the value found in the characteristic solution of 0.193. This indicates that the characteristic method of solution for this particular type of problem gives very accurate results.

Equation (24) can be generalized to arbitrary α, β very easily.

APPENDIX I

By eliminating all unknowns but one, we solve for σ^*, ϵ^* in (4) to find that they satisfy the same partial differential equation

$$\begin{aligned}&\beta\sigma_{\tau\tau\tau}{}^*-\sigma_{\xi\xi\tau}{}^*+(1+\alpha+\beta)\sigma_{\tau\tau}{}^*-\sigma_{\xi\xi}{}^*+\alpha\sigma_\tau{}^*=0\\&\beta\epsilon_{\tau\tau\tau}{}^*-\epsilon_{\xi\xi\tau}{}^*+(1+\alpha+\beta)\epsilon_{\tau\tau}{}^*-\epsilon_{\xi\xi}{}^*+\alpha\epsilon_\tau{}^*=0.\end{aligned} \quad (A.1)$$

We also have from (4) that

$$\sigma_\xi{}^*=v_\tau{}^*, \quad (A.2)$$

and hence from (A.1)

$$\beta v_{\tau\tau\tau\tau}{}^*-v_{\xi\xi\tau\tau}{}^*+(1+\alpha+\beta)v_{\tau\tau\tau}{}^*+v_{\xi\xi\tau}{}^*+\alpha v_{\tau\tau}{}^*=0. \quad (A.3)$$

Integrating with respect to τ gives

$$\beta v_{\tau\tau\tau}{}^*-v_{\xi\xi\tau}{}^*+(1+\alpha+\beta)v_{\tau\tau}{}^*-v_{\xi\xi}{}^*+\alpha v_\tau{}^*=0, \quad (A.4)$$

where we have used the conditions that the body is initially unstressed and unstrained with zero velocity.

The fact that all of the dependent variables satisfy the same differential equation gives us much additional information when a single characteristic solution is found. Because of the range of solutions to be investigated in problems with two variable parameters any additional information without computing another solution is highly desirable.

APPENDIX II

We use the Laplace transform method and for simplicity will consider the case of $\alpha=\beta=1$. The general α, β case can be considered similarly.

We apply the Laplace transform

$$\bar{\varphi}=L\{\varphi\}=\int_0^\infty \exp(-p\tau)\varphi(\tau)d\tau \tag{B.1}$$

to the set of Eq. (6) and obtain after simplification

$$\bar{\sigma}_{\xi\xi}{}^*-\left(\frac{p^2}{1+p}+p+p^2\right)\bar{\sigma}^*=0 \tag{B.2}$$

$$\bar{\sigma}_\xi{}^*=p\bar{v}^*. \tag{B.3}$$

We have used the initial conditions that

$$\sigma^*(\xi,0)=\epsilon_2{}^*(\xi,0)=\epsilon^*(\xi,0)=v^*(\xi,0)=0. \tag{B.4}$$

For boundary conditions we have that

$$\begin{aligned} v^*(0,\tau)&=-g_0 \\ \lim_{\xi\to\infty}\sigma^*(\xi,\tau)&=0. \end{aligned} \tag{B.5}$$

These boundary conditions transform to

$$\begin{aligned} \bar{v}^*(0,p)&=-\frac{g_0}{p} \\ \lim_{\xi\to\infty}\bar{\sigma}^*(\xi,p)&=0. \end{aligned} \tag{B.6}$$

Solving (B.2), (B.3) using (B.6) we have

$$\bar{\sigma}^*=g_0\exp-\left(\frac{p^2}{1+p}+p+p^2\right)^{\frac{1}{2}}\xi\Big/\left(\frac{p^2}{1+p}+p+p^2\right)^{\frac{1}{2}} \tag{B.7}$$

$$\bar{v}^*=-g_0\exp-\left(\frac{p^2}{1+p}+p+p^2\right)^{\frac{1}{2}}\xi. \tag{B.8}$$

From (B.7) we find that for the end of the bar

$$\bar{\sigma}^*(0,p)=\frac{g_0(1+p)^{\frac{1}{2}}}{(p)^{\frac{1}{2}}(p+a)^{\frac{1}{2}}(p+b)^{\frac{1}{2}}}, \tag{B.9}$$

where

$$a=\frac{3}{2}-\frac{(5)^{\frac{1}{2}}}{2},\quad b=\frac{3}{2}+\frac{(5)^{\frac{1}{2}}}{2}.$$

Formally, the inversion integral gives

$$I=\sigma^*(0,\tau)=\frac{g_0}{2\pi i}\int_{\gamma-i\infty}^{\gamma+i\infty}\frac{e^{\lambda\tau}(\lambda+1)^{\frac{1}{2}}}{(\lambda)^{\frac{1}{2}}(\lambda+a)^{\frac{1}{2}}(\lambda+b)^{\frac{1}{2}}}d\lambda, \tag{B.10}$$

where we have $-\pi<\arg\lambda\leq\pi$. It is easily seen that

$$I=e^{-t}\frac{\partial}{\partial t}(e^tI'), \tag{B.11}$$

with

$$I'=\frac{g_0}{2\pi i}\int_{\gamma-i\infty}^{\gamma+i\infty}\frac{e^{\lambda\tau}}{(\lambda)^{\frac{1}{2}}(\lambda+1)^{\frac{1}{2}}(\lambda+a)^{\frac{1}{2}}(\lambda+b)^{\frac{1}{2}}}d\lambda. \tag{B.12}$$

From Churchill,[7]

$$L\left\{e^{-\frac{1}{2}(a+b)t}I_0\left(\frac{a-b}{2}t\right)\right\}=\frac{1}{(p+a)^{\frac{1}{2}}(p+b)^{\frac{1}{2}}}, \tag{B.13}$$

and consequently we have from the convolution theorem that

$$I'=g_0\exp(-t/2)\int_0^t e^{-\lambda}I_0\left(\frac{(5)^{\frac{1}{2}}}{2}\lambda\right)\times I_0[\tfrac{1}{2}(t-\lambda)]d\lambda. \tag{B.14}$$

Hence,

$$\sigma^*(0,\tau)=\exp(-3\tau/2)I_0\left(\frac{(5)^{\frac{1}{2}}}{2}\tau\right)+\frac{1}{2}\int_0^\tau\exp(-\lambda-\tau/2)I_0\left(\frac{(5)^{\frac{1}{2}}}{2}\lambda\right)\times\{I_0[\tfrac{1}{2}(\tau-\lambda)]+I_1[\tfrac{1}{2}(\tau-\lambda)]\}d\lambda. \tag{B.15}$$

[7] R. V. Churchill, *Modern Operational Mathematics in Engineering* (McGraw-Hill Book Company, Inc., New York, 1944), p. 297.

Thermionic Emission and Electron Diffraction from Thin Films of Barium Oxide*

PAUL N. RUSSELL† AND A. S. EISENSTEIN
Department of Physics, University of Missouri, Columbia, Missouri
(Received September 1, 1953)

A high-vacuum, sealed-off glass, electron diffraction tube was developed for observing the physical structure of evaporated BaO films on a nickel substrate. Simultaneously, measurements of thermionic emission could be made. The thermionic activity of the deposited film increased until approximately 20 monolayers (measured using a radioactive tracer technique) had been laid down, and no further increase was found up to 50 monolayers. This emission density (0.4A/cm² at 1000°K) is comparable with that obtainable from sprayed oxide coatings. With the receiver at less than 450°K the oxide film showed an amorphous structure. With the receiver at 800°K the film was crystalline and was believed to exist in clumps. This crystalline structure became amorphous if the films were heated to 1070°K and this form persisted as the oxide film was removed by evaporation. The relation of the thermionic emission from thin BaO films to that from sprayed oxide cathodes is discussed.

INTRODUCTION

ONE salient feature of all previous studies[1–3] of the thermionic emission capabilities of monolayer films of evaporated BaO is the close similarity between both the magnitude and the temperature dependence of this emission and that of thick, sprayed oxide-coated cathodes. This similarity has given rise to speculation that the emission mechanism operating in these two physically different structures may in fact be similar.

Davisson and Pidgeon,[1] who first studied the electron emission from thin BaO films found that optimum emission occurred for a coverage of about one monolayer. However, their attempt to interpret the details of the emission *vs* thickness characteristic was based on a model in which the oxide was not uniformly distributed but was laid down in clumps. Moore and Allison[3] deposited by evaporation monomolecular films of BaO and SrO onto tungsten and molybdenum surfaces and observed a monotonic increase of emission with thickness through twenty-five molecular layers. The work function of these surfaces was found to decrease rapidly for. the first one-tenth monolayer and then only gradually, reaching a limiting value of about 1 ev after a few monolayers were deposited. An emission current of nearly 1 A/cm² at 1000°K could be obtained from these monomolecular films. A model consisting of a uniform distribution of oriented molecular dipoles was assumed and was used to explain the observed emission phenomena.

X-ray diffraction techniques have been used extensively in studies of the physical structure of the oxide cathode coating and the interface region. Because of its reduced penetration and greater scattering efficiency, a beam of electrons diffracted from a sample should provide more information about the nature of the surface layers than is obtainable from x-ray diffraction studies. Nevertheless, electron diffraction from oxide cathode surfaces has contributed little new information beyond what has been obtained with x-ray scattering.

The vacuum conditions found in most electron diffraction units are so poor that, although the oxide sample is usually prepared by conversion from the carbonate within the unit, the samples are hardly typical oxide coated cathodes. Only Darbyshire[4] reports finding BaO on the surface of cathodes converted from $BaCO_3$. Other studies[5,6] produced patterns of $Ba(OH)_2$ and BaS which may have resulted from contamination of the oxide surface by a reaction with residual vapors. Whereas most low-energy electron diffraction studies are carried out in thoroughly evacuated, sealed-off tubes, this technique does not seem to have been used extensively for high-energy electron diffraction work. Darbyshire[7] has described a sealed-off, diffraction tube built for demonstration purposes. This resembled a cathode-ray tube, and the diffraction pattern was viewed on a phosphor screen.

As a preliminary to the work reported here an investigation[8] was made to compare various means for detecting and recording electron diffraction information that did not involve the use of photographic plates within the diffraction chamber. The technique of allowing the diffraction patterns to fall on a phosphor screen within the evacuated chamber and recording the pattern photographically with a camera external to the vacuum was found to be quite satisfactory. Since quantitative measurements of diffraction intensities are seldom required, a calibration of the system is not necessary.

A study of the thermionic emission characteristics of thin BaO films was undertaken in a high-vacuum, sealed-off, electron diffraction tube in which observa-

* Supported in part by the U. S. Office Naval Research.
† Now at General Electric Company, Syracuse, New York.
[1] Davisson and Pidgeon, Phys. Rev. **15**, 553 (1920).
[2] A. W. Hull, Phys. Rev. **56**, 86 (1939).
[3] G. E. Moore and H. W. Allison, Phys. Rev. **77**, 246 (1950).
[4] J. A. Darbyshire, Proc. Phys. Soc. (London) **50**, 635, 964 (1938).
[5] H. Gaertner, Phil. Mag. **19**, 82 (1935).
[6] Huber and Wagener, Z. tech. Phys. **23**, 1 (1942).
[7] J. A. Darbyshire, Elec. Eng. **22**, 385 (1950).
[8] A. Eisenstein, paper presented at ASXRED summer meeting, 1948.

tions of the physical structure of the film could be correlated with measurements of the electron emission. Changes of both structure and emission accompanying heat treatment of the sample could be followed as the BaO film was removed by evaporation.

EXPERIMENTAL PROCEDURES

I. Determination of Evaporation Rate

Although the vapor pressure of BaO is well known,[3] the rate of deposition was verified experimentally using a structure identical with that to be used in the final diffraction tube. This preliminary experiment was carried out to avoid possible error in the calculated amount of deposited BaO which might result from a temperature gradient in the evaporator or from an inaccurate calibration of the monitoring thermocouple. At the temperature used for evaporation, 1310°K, the rate of deposition doubles for a temperature change of about 15°.

The deposition rate was determined by the use of radioactive tracer techniques in an evacuated tube. $Ba^{140}CO_3$ ‡ was applied to a platinum§ evaporator structure identical with that used in the diffraction tube. Above the evaporator at a distance of 7 mm was mounted a plane nickel shield with a 5-mm hole punched in its center. Above the hole was suspended one of several nickel sample buttons which could be manipulated into position for a given time of evaporation and could then be removed, making way for another. During a period of evaporation the temperature of the evaporator was held constant at 1310°K as measured by a Pt−Pt+10-percent Rh thermocouple. When each sample had been exposed to the molecular beam for a definite time ranging from 2 minutes to 10 hours, the tube was cracked open and the radioactivity of each sample was compared with that of a standard aliquot of the original $BaCO_3$. The plot of Geiger counts per minute *vs* evaporation time on a log-log scale gave a straight line with a slope of unity, as shown by the upper curve in Fig. 1.

Closer examination of the data revealed that the count due to the radioactivity of the first button, which had received a 20-minute deposition, was about 30 percent greater than the count of the last button which had also received a 20-minute evaporation, but 19 hours later. Ba^{140} with a half-life of 13.4 days decays into La^{140} with a half-life of 40.2 hours. In the aliquot and on the evaporator the lanthanum is in transient equilibrium with the barium, but due to the refractory nature of La_2O_3 it is reasonable to assume that no significant amount of this oxide is evaporated at 1310°K, only BaO. The lower curve of Fig. 1 results

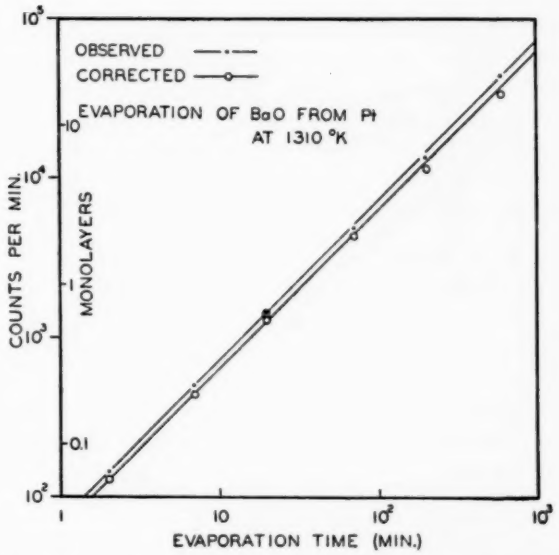


FIG. 1. Counting rate produced by $Ba^{140}O$ reaching the collector plate as a function of the evaporation time at 1310°K. Lower curve corrected for activity of daughter La^{140}.

from correcting the observed counting rates and brings the two 20-minute depositions to within 7 percent. Furthermore, the corrected deposition rates are found to agree within a factor of two with the rates computed from the vapor pressure of BaO and geometrical considerations. For convenience the term "monolayer" will be used to represent an average surface density of 10^{15} BaO molecules/cm², whether the molecules form a true monolayer or aggregate at isolated regions of the surface. Figure 1 shows that one monolayer is deposited in 27 minutes at 1310°K.

II. Diffraction Tube Construction and Operation

Preliminary studies of thermionic emission from evaporated BaO films were carried out in evaporator-diode tubes having the same geometry as that used in the electron diffraction tubes. The results presented here are based on the performance of five evaporator-diode tubes and four electron diffraction tubes. Only the electron diffraction tube will be described in detail.

This tube, shown in Fig. 2, consisted of an elongated cathode-ray tube blank with a commercial electron gun of the type used in the 10FP4 tube‖ at one end and a phosphor screen of the P4 variety in the other end. The evaporator-diode structure was placed in a cross arm midway between the gun and the screen, in such a position that the polished nickel sample would intercept the electron beam at near grazing incidence.

The evaporator-diode structure consisted of a polished receiver of pure electrolytic nickel¶ indirectly heated by a tungsten coil. The temperature was meas-

‡ Prepared from Ultra Pure Barium Carbonate, Mallinckrodt Chemical Works, St. Louis, Missouri.

§ 99.9-percent pure Pt obtained from J. T. Baker Co., Newark, New Jersey. Evaporator structure outgassed at 1500° K in vacuum before applying the $BaCO_3$.

‖ These guns were kindly supplied by J. J. O'Callaghan, Rauland Corporation, Chicago, Illinois.

¶ 1001 Nickel obtained from E. M. Wise, International Nickel Company, New York, New York.

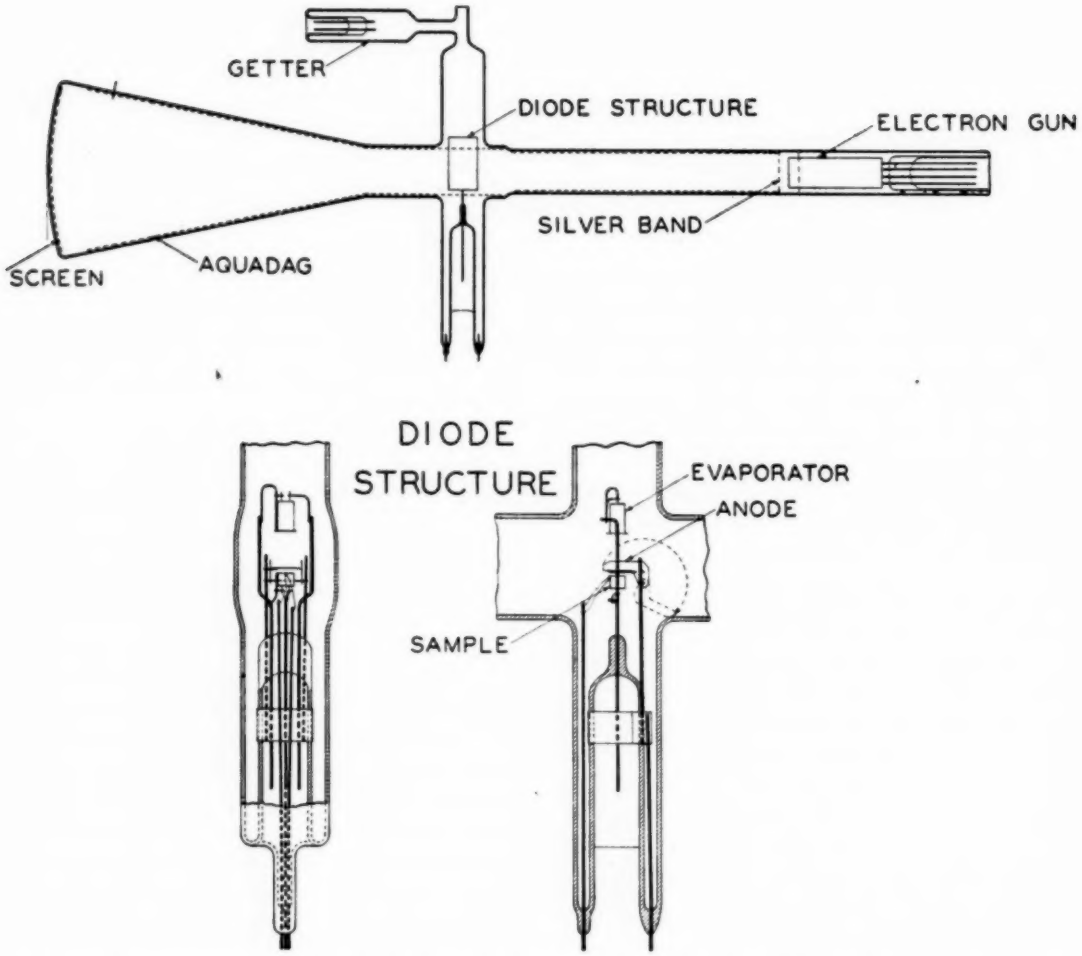


FIG. 2. Complete electron diffraction tube showing electron gun, evaporator-diode structure, and phosphor viewing screen.

ured by a Mo-Ni thermocouple. Immediately above at a distance of 17 mm was mounted the platinum faced evaporator. This evaporator was indirectly heated and its temperature was measured by a Pt−Pt+10-percent Rh thermocouple. Between the evaporator and the receiver was located a hinged tantalum anode which could be rotated out of the way during evaporation of the BaO and also during the observation of the electron diffraction pattern; but which in its closed position served as a shield for the clean nickel receiver surface during conversion of the $BaCO_3$ to BaO, and as an anode for the measurement of thermionic emission from the thin deposit of BaO.

The metal components of this structure were individually vacuum degassed or hydrogen fired before assembly. The platinum evaporator was placed behind a mask which had a 0.1-cm² aperture, and was sprayed with a water suspension of Ultra Pure $BaCO_3$ to a surface density of 15 mg/cm². The structure was immediately assembled and sealed onto the electron diffraction tube. The tube was then evacuated, processed, and sealed off at a pressure of about 5×10^{-8} mm of Hg, and the Batalum getter flashed. It was then "aged" by applying successively higher voltages to the accelerating electrodes until it would operate stably at well over 25 kv. During this process the pressure as measured by an attached Bayard-Alpert ionization gage did not rise perceptibly. The pressure after aging was less than 1×10^{-8} mm of Hg.

The electron beam was brought to a focus on the phosphor screen by means of a commercial electromagnetic focusing coil. Since the sample was rigidly mounted in the tube, it was necessary to deflect doubly the electron beam to the sample at grazing incidence. The resulting diffraction pattern was photographed externally by means of a Dumont 35-mm oscilloscope camera, $f/3.5$, using Eastman Super XX film. Exposures of 2, 4, and 8 minutes were made of each pattern, and the film was developed before the physical state of the sample was changed. Halation rings, usually present because of the primary beam of high-energy electrons impinging on the phosphor and producing light which is partially reflected internally in the glass face plate, were eliminated by removing

the phosphor from a small area of the screen in that region during the construction of the tube. The film was then viewed, and the radii of the rings measured with a microprojector. Calibration of the over-all optical magnification was accomplished by photographing an opaque patch of known dimensions which had been secured to the outer surface of the face plate.

III. Electron Emission Measurements

Thermionic emission currents from the cathode at a given temperature were measured for anode potentials from 20 volts positive to several volts negative and were used to construct retarding potential plots from which the zero field emission current densities J_0 were obtained. This procedure was repeated for a series of cathode temperatures measured in descending order. A Richardson plot representing the particular state of activation of the cathode was made and used to evaluate the work function and the characteristic emission of the cathode at 1000°K.

EXPERIMENTAL RESULTS

In its initial state, the thermionic emission of the "clean" nickel was too low to be measured with accuracy with the available equipment; it was of the order of 10^{-12} amp/cm² at 1000°K. The receiver was held at 800°K while one monolayer of BaO was deposited. The subsequent temperature dependence of emission is shown by curve *A* of Fig. 3. When the sample was heat treated for 10 minutes at 1070°K the emission characteristic fell to curve *B*. Further heat treatment of the sample for 10 minutes at 1170°K gave curve *C*. In this state the electron diffraction pattern of the surface corresponded to that of nickel rather than BaO, and further heating did not further reduce the emission. The subsequent curves in Fig. 3 represent two complete repetitions of this cycle. The close similarity of these three sets of curves indicates the reproducibility of the electron emission characteristics of deposited BaO films on a nickel substrate. When this procedure was repeated using 3 monolayer depositions, longer heat treatments of the sample were necessary to reduce the emission, but in no case was it reduced below that of curve *C*.

More than 50 sets of photographs were made of the diffraction patterns produced in this tube representing different physical states of the emitting surface. For brevity and convenience all diffraction patterns observed in this study may be classified into 4 types and combinations thereof: (I) "Laue type" pattern; (II) broad ring pattern, some rings exhibiting preferred orientation; (III) sharp ring pattern; and (IV) diffuse pattern. These are shown in Fig. 4.

Type I was observed only from the surface of "clean" nickel either before deposition of BaO, or after a heat treatment sufficiently severe to cause it to re-evaporate. Photo micrographs and electron micrographs, using the replica technique, were taken of the polished nickel surfaces. Crystals having an average dimension of 0.025 mm are present in the nickel substrate; hence, as many as 10^4 crystals are bathed in a 1-mm² beam of electrons grazing the surface. Of this number perhaps 20 are properly oriented to satisfy the Laue conditions for diffraction and produce the spots observed in this type of pattern. The presence of this type of pattern and the low emission represented by the Richardson plot Fig. 3, curve *C*, were the criteria for a "clean" surface following the re-evaporation of the BaO.

The Type II pattern was observed after the deposition of one or more monolayers of BaO onto a receiver held at 800°K. Since it is generally believed that electron diffraction patterns are not observable from crystallites having less than five or ten diffracting planes, this deposit is probably not a true monolayer. Rather, it is a "patchy" deposit in which only portions of the surface are covered by the BaO and those portions are covered to a depth of several molecular layers. It will be remembered that one deposited monolayer was defined as 10^{15} molecules per cm² irrespective of its physical arrangement. Measurement of the radii of the diffraction rings indicated that in every case the deposited material had a face centered

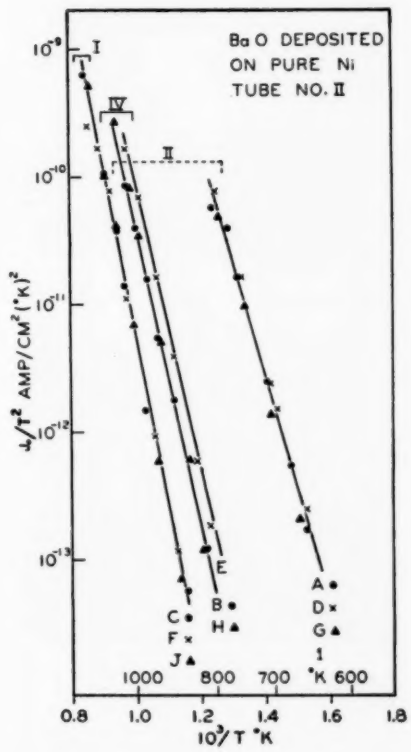


FIG. 3. Typical Richardson plots of the thermionic emission from BaO deposited on a pure nickel base receiver. (A) One monolayer deposition; (B) after heat treatment for 10 minutes at 1070°K; and (C) after 10 minutes at 1170°K. [Process repeated showing reproducibility, in sequence (D), (E), (F), (G), (H), and (J). Type of diffraction pattern observed shown at top.]

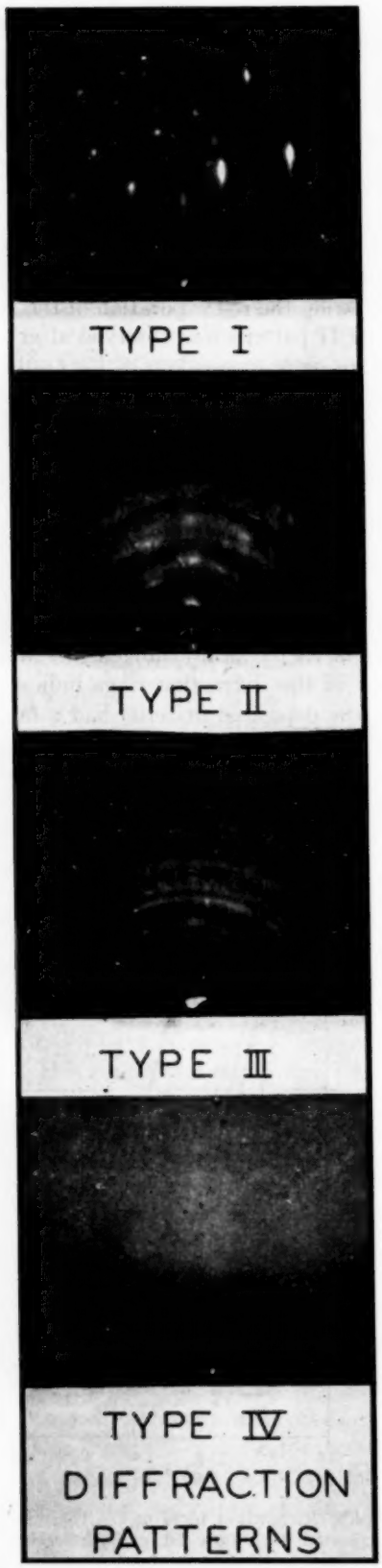


FIG. 4. Four types of electron diffraction patterns observed: (I) "Laue type" pattern of the nickel receiver; (II) broad ring pattern of BaO showing some preferred orientation; (III) sharp ring pattern of BaO; and (IV) diffuse pattern of BaO.

cubic structure with a lattice constant of 5±1A. Barium oxide is face centered cubic with a lattice constant of 5.53A. For thin deposits the resulting pattern showed variation of intensity along a given ring. This was attributed to a preferred orientation of the deposited crystallites with the (1,0,0) planes of the BaO coinciding with the nickel surface. Patterns from thicker deposits showed less preferred orientation, indicating that the orientation of the upper layers does not necessarily correspond to that of the first deposited layers and may thus represent the growth of new crystallites. The breadth of the diffraction rings indicates that the crystallites were very small, probably less than 100A in size, or that the lattice was rather imperfectly developed, or both.

The type III pattern was similar to type II, but differed from it in that the rings were quite sharp, indicating crystallites larger than 100A in size and having no preferred orientation. This pattern was not observed when BaO was deposited on the pure nickel substrate, even for deposits of up to 50 monolayers. The sharp ring pattern was found only in some subsequent depositions on nickel alloy substrates and is shown here as evidence that thick films of BaO having well-developed crystals may be obtained by evaporation and detected by electron diffraction.

The very diffuse type IV pattern was observed under two conditions, (1) when BaO was deposited on a receiver held at a temperature less than 450°K, and (2) as a result of very high temperature heat treatments of a crystalline BaO film. This diffuse scattering of the impinging electrons is probably produced by an amorphous layer of BaO on the metallic substrate. If the temperature is sufficiently low that surface migration does not occur, a regular lattice array of the deposited molecular dipoles is rather improbable. Subsequent layers on this substrate likewise may have a random arrangement and an amorphous structure would be formed. Electrons scattered from such a random arrangement of molecules would produce this diffuse pattern.

Combinations of these patterns also appeared. Ordinarily, BaO deposited upon the base metal gave a ring-type pattern. With continued heat treatment of the sample the type II pattern seemed to fade into type IV, which in turn changed into type I. The Roman numerals bracketing the curves of Fig. 3 indicate the type of pattern which was observed for each state of the emitting surface.

The emission obtained from successive depositions of BaO is shown in the Richardson plots of Fig. 5. During this deposition the receiver was held at 800°K. It will be noted that the thermionic activity rises until about 27 monolayers are deposited, then there is no significant change up to the maximum amount of 51 monolayers. At this maximum activity the emission was 0.4 A/cm² at 1000°K with a work function of about 1.3 ev. The type II diffraction pattern was found in each case.

After this thick coating had been deposited, an attempt was made to remove it by heat treatment and to study the subsequent decrease in thermionic activity. These results are summarized in Fig. 6. Curve λ was obtained after heating the sample at 1170°K for 20 minutes. Curve μ which shows a tenfold decrease in activity below curve λ was obtained after heating 180 minutes at 1170°K and 80 minutes at 1270°K. From Richardson plots taken at intervals during this treatment it was evident that further heating at 1170°K had little effect. Curve φ represents the lowest state of activity obtained in this sequence, with an emission of 5×10^{-4} amp/cm² at 1000°K and a work function of 2.6 ev. Further reduction of activity was not attempted, since the type I diffraction pattern characteristic of the nickel substrate had replaced the type IV diffuse pattern. Table I summarizes the results obtained in these two experiments on deposition and re-evaporation. Listed are the amount of deposition, characteristic emission at 1000°K, work function, subsequent heat treatments of the sample, and types of diffraction patterns observed.

Similar observations made in four diffraction tubes gave the following results:

(1) Evaporated BaO films remained stable over periods of weeks with respect to their thermionic emission characteristics and their physical structure.

(2) The deposited BaO assumed different physical forms depending upon the temperature of the receiver at the time of deposit. An amorphous state, giving a type IV pattern, was produced when the receiver was below 450°K. Crystalline aggregates, giving a type II pattern, were formed at 800°K. The amorphous state changed to the crystalline state when the sample was held at 800°K for a short time.

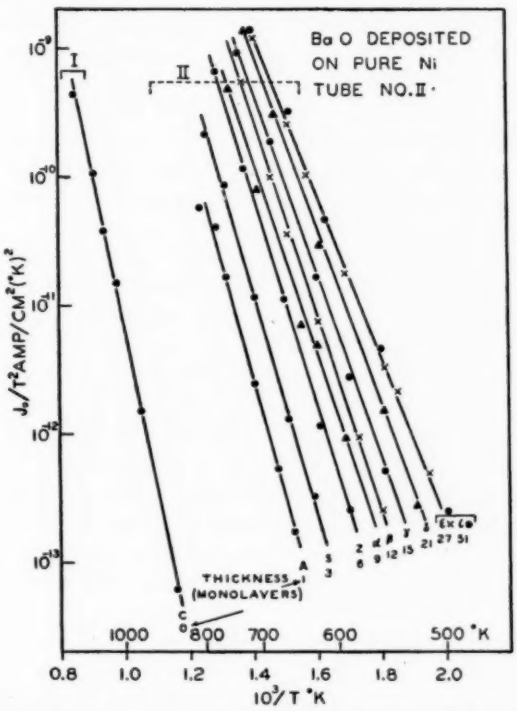


FIG. 5. Richardson plots of thermionic emission from successive depositions of BaO, 1 to 51 monolayers. Type of diffraction pattern observed shown at the top.

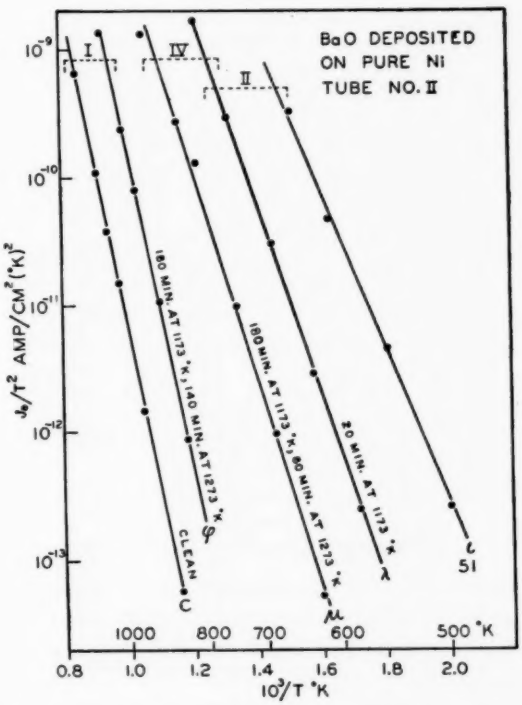


FIG. 6. Richardson plots of thermionic emission from deposited BaO layer accompanying removal by re-evaporation. Type of diffraction pattern observed shown at the top.

(3) Preferred orientation in the diffraction patterns of a few monolayers deposited at 800°K gave evidence that the (1,0,0) planes of the crystallites coincide with the surface of the underlying substrate.

(4) Maximum thermionic emission was reached at about 20 monolayers of BaO with little further change up to 50 monolayers. This emission level is comparable with that obtainable from sprayed oxide coatings.

(5) Subsequent heating of the crystalline films to 1070°K or above caused them to become amorphous. This amorphous form persisted as the oxide was removed by evaporation and was accompanied by a decreasing thermionic emission.

(6) Continued re-evaporation of BaO from the sample reduced the emission to a common minimum level which was six orders of magnitude above the emission obtainable from clean nickel at 1000°K.

(7) It was possible to observe the same thermionic emission characteristics, work function, and characteristic emission at 1000°K, from samples in two different physical states as judged by the electron diffraction patterns. For example, a thin crystalline deposit and an amorphous film of unknown thickness which was formed by heat treatment of a thick deposit gave almost identical Richardson plots. Thus, a given physical state of the oxide is not uniquely defined by its thermionic emission characteristics.

TABLE I.

State of activation	Work function (ev)	Characteristic emission (A/cm² at 1000°K)	Monolayers deposited	Diffraction pattern	Total heat treatment
C	2.66	7.0×10^{-6}	0	I	
A	1.90	1.8×10^{-2}	1	II	
S	1.81	6.0×10^{-2}	3	II	
Z	1.66	2.0×10^{-1}	6	II	
α	1.62	2.0×10^{-1}	9	II	
β	1.53	4.0×10^{-1}	12	II	
γ	1.42	4.0×10^{-1}	15	II	
δ	1.39	4.0×10^{-1}	21	II	
ϵ	1.27	4.0×10^{-1}	27	II	
ι	1.37	4.0×10^{-1}	51	II	
λ	1.53	1.6×10^{-1}		II,IV	10 min 1170°K
μ	1.82	2.0×10^{-2}		IV	180 min 1170°K 80 min 1270°K
φ	2.58	5.0×10^{-4}		I	180 min 1170°K 140 min 1270°K

INTERPRETATION OF RESULTS

Several features of the observed changes in thermionic emission and physical structure of BaO films accompanying deposition and heat treatment were unexpected and require interpretation.

Films laid down on a receiver at or below 450°K showed an amorphous structure but became crystalline when heated to 800°K whereas films laid down on a receiver at 800°K were crystalline at room temperature. These observations seem to indicate that the surface and bulk mobility of BaO is so low at 450°K that the molecules cannot orient themselves properly to build up the crystal lattice. Thus they have a random arrangement in the film. At 800°K the diffusion rate has increased sufficiently to permit the rearrangement of molecules into a crystal lattice, which form represents the state of lowest free energy at that temperature. This argument may also be used to explain the return to the amorphous state in samples heated to 1070°K and above. It is possible that at elevated temperatures the free energy is minimized by increasing the disorder (entropy) as would happen in a change from the crystalline to the amorphous form, but at best this is only a speculation. Since all diffraction patterns were taken with the sample at near room temperature, it must be concluded that the physical state of the system was "frozen in" as the sample was cooled from the elevated temperatures.

The observation that maximum emission is obtained only when some 20 monolayers of BaO have been deposited is unexpected although not in disagreement with the results of Moore and Allison. However, metallic monolayers yield optimum emission for a surface coverage of from 0.67 to 1.0 monolayers.[9] Since a type II electron diffraction pattern was observed from a deposition of one monolayer of BaO, it may be reasonably assumed that this material existed in clumps or patches. Within each clump there must be a sufficient number of molecules to comprise a fairly well-ordered crystal lattice to produce the observed diffraction. The thermionic emission current density J_0 which was measured with a low applied field from this patchy surface should be considerably less than the algebraic sum of the true zero field current densities from each patch. This reduced emission from the oxide clumps is a result of the complex electric field established by the surrounding areas of high work function.[10] The observed current density from this patchy monolayer was found to be 1.8×10^{-2} A/cm² at 1000°K, only five percent of the emission that could be obtained from films of greater thickness. As the amount of deposited material increases, the fraction of the surface covered by these emitting patches likewise increases until complete coverage is obtained at about 20 deposited monolayers. Further depositions cause no further increase in emission above this level.

The observation that continued re-evaporation of BaO from the sample reduced the emission to a common minimum level, six orders of magnitude above the emission from clean nickel at 1000°K, must be interpreted in terms of a residue of BaO or some BaO−Ni compound. Electron diffraction patterns of the surface preclude the existence of these materials there. It seems likely that some of the BaO may have migrated into the crevasses of the grain boundaries which were clearly seen in photomicrographs of the surface. Van der Waal's forces[11] would be large in these regions and tend to prevent migration to the surface even at temperatures which would cause rapid evaporation from the surface. A slow migration from these interstices to the surface could produce this constant level of emission which is observed over a long period of time.

Thin evaporated films of BaO on Ni give a characteristic emission at 1000°K of about 0.4A/cm² with a

[9] I. Langmuir and J. B. Taylor, Phys. Rev. 44, 423 (1933).

[10] R. R. Brattain and J. A. Becker, Phys. Rev. 43, 428 (1933).

[11] J. H. de Boer, *Electron Emission and Adsorption Phenomena* (MacMillan Company, New York, 1935), Chapter 2.

work function of between 1.2 and 1.3 ev, both values comparing favorably with the reported[3] emission values for thick, sprayed oxide cathodes. It is now worth considering whether the emission mechanisms operating in these two physically different structures are related. We may also consider whether the emission from a thin BaO film on the cathode base metal plays any essential role in the normal operation of thick, sprayed oxide cathodes.

If the conduction of electrons from the base metal to the external vacuum surface in sprayed cathodes takes place by means of crystal conduction[12,13] and is followed by thermionic emission from the vacuum surface, it is difficult to assign any essential role to a thin oxide film present at the base metal-coating interface. However, the similarity in emission from a film composed of crystallites 20 monolayers thick and from the micron-size crystals which comprise the surface of sprayed oxide coatings is not too surprising. These small crystallites have a fairly well-developed crystal structure, as evidenced by the electron diffraction, and thus may be expected to exhibit electronic properties similar to the larger crystals.

If, on the other hand, at normal operating temperatures an electron gas within the pores[14] of the sprayed coating conducts electrons through the coating, an emission mechanism at the base metal-coating interface is desired. A thin film of the oxide on the base metal could supply the necessary emission to provide an adequate electron gas density. This density would be maintained in the pores by absorption and re-emission from the pore walls. Since the density of the electron gas at the vacuum surface of the cathode is primarily a function of the absorption and re-emission from the oxide crystals in that region, the thin oxide film at the base metal serves only to maintain an adequate density of the electron gas at the interface. It is difficult to believe that during cathode processing at least a few monolayers of the oxide have not been evaporated onto the base metal from the adjacent oxide crystals. This pore conduction hypothesis has been given added impetus by the recent work of Hensley[15] and Young.[16] Although in this cathode model the oxide film at the base metal serves an important role, the emission characteristics of the cathode are determined by the oxide near the surface. Therefore, the similarity in cathode emission with that of a thin film must be interpreted as in the previous paragraph.

This paper summarizes the results of a study of the thermionic emitting surface BaO on pure nickel. Somewhat similar results were obtained when BaO films were evaporated onto a 4.7-percent W—Ni alloy. In general, the results were less reproducible with the alloy base metal, but the levels of attainable emission were the same as on the pure nickel. No evidence of an interface compound was found from the diffraction patterns. It is anticipated that the sealed-off, glass electron diffraction tube developed in this study will find application in the examination of other surfaces which must be handled under very high vacuum conditions.

[12] Hannay, MacNair, and White, J. Appl. Phys. **20**, 669 (1949).
[13] D. A. Wright, Phys. Rev. **82**, 574 (1951).
[14] R. Loosjes and H. J. Vink, Philips Research Repts. **4**, 449 (1949).
[15] E. B. Hensley, J. Appl. Phys. **23**, 1122 (1952).
[16] J. R. Young, J. Appl. Phys. **23**, 1129 (1952).

JOURNAL OF APPLIED PHYSICS VOLUME 25, NUMBER 8 AUGUST, 1954

On the Nature of Radiation Damage in Metals

JOHN A. BRINKMAN
North American Aviation, Inc., Downey, California
(Received October 5, 1953)

The nature of the permanent damage retained in metals from irradiation has been investigated in somewhat greater detail than has been done in the past. The usual assumption has been that the damage in all metals consists chiefly of interstitial-vacancy pairs. The model presented in this paper reduces to this picture for the light elements but introduces a new concept in the case of damage in the heavy metals, called a displacement spike. Calculations are made from which one can estimate the relationship between the density of interstitial-vacancy pairs and the temperature of the associated thermal spike. An assumption regarding the extent to which interstitial-vacancy pairs persist throughout the duration of the thermal spike has been made, based upon these calculations. The number of interstitial-vacancy pairs predicted in the heavy elements is considerably smaller than that predicted by the former model. A mechanism is proposed by which small dislocation loops can be produced in the heavier metals by irradiation.

This article is based upon studies conducted for the U. S. Atomic Energy Commission under Contract AT-11-1-GEN-8.

INTRODUCTION

CHARGED particle or neutron irradiation is known to produce lattice changes in metals which can be retained as permanent damage as long as the metals are held at sufficiently low temperature. The nature of these lattice distortions has been the subject of considerable theoretical study, in both the open and classified literature, by such workers as F. Seitz, H. Brooks,

J. D. Ozeroff, H. M. James, F. W. Brown, and M. M. Mills. Seitz has described the development and present status of the theory concerning such radiation induced lattice imperfections in solids.[1] For metals, according to the presently accepted picture, the damage is comprised of two aspects (1) Frenkel defects, or interstitial-vacancy pairs and (2) effects resulting from thermal spikes.

The assumption has frequently been made[1] that essentially every atom which receives an energy greater than a certain threshold necessary to displace it from its lattice site will persist as a permanent interstitial atom, while its lattice site will remain as a vacancy. The number of Frenkel defects which should be produced in metals by irradiation has been calculated by Seitz on the basis of this assumption.[2] Thermal spikes have been assumed capable of producing effects of similar nature to those resulting from heating and rapidly quenching the metal. The disordering of ordered $AuCu_3$ by neutron irradiation observed by Siegel[3] has been cited as an example of such an effect.[1,2] The Frenkel defects produced by a given primary knock-on atom will all lie within the limits of the associated thermal spike. The above assumption regarding the persistence of the radiation induced interstitials and vacancies therefore must include the more basic assumption that the time duration of the thermal spike is too short to permit appreciable annealing of the Frenkel defects.

The present paper represents an attempt by the author to construct a model of radiation damage in metals based upon somewhat different assumptions. While a detailed theoretical investigation of these assumptions, as well as those made by other workers, would be desirable, the author can at present only give qualitative arguments for them and appeal to experimental work to differentiate between the two models.

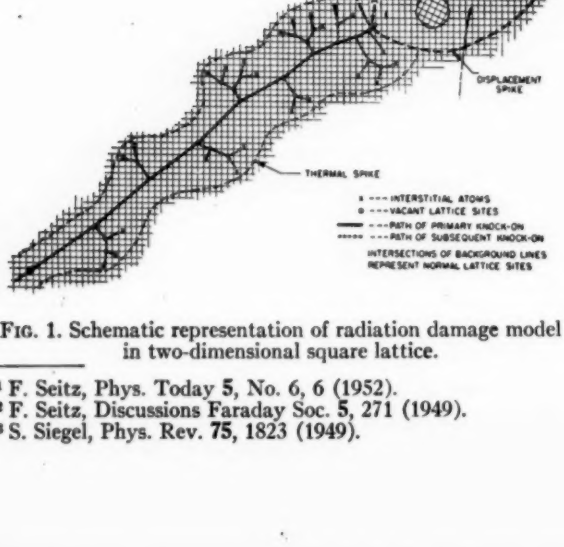


FIG. 1. Schematic representation of radiation damage model in two-dimensional square lattice.

[1] F. Seitz, Phys. Today **5**, No. 6, 6 (1952).
[2] F. Seitz, Discussions Faraday Soc. **5**, 271 (1949).
[3] S. Siegel, Phys. Rev. **75**, 1823 (1949).

RADIATION DAMAGE MODEL

In this section, the atomistic picture of radiation damage which will be developed later is described in a qualitative manner along with the assumptions upon which it is based. These are contrasted with the assumptions made by others, and the resulting models are compared.

The present paper deals only with the damage produced in a metal by knock-on atoms and the dependence of this damage on the energy of the knock-ons. This treatment is therefore independent of the type of bombarding particle used. To apply the general concepts developed in the present paper to any specific type of irradiation, one must calculate the energy spectrum of the primary knock-ons resulting from such irradiation. By this means, one should obtain a more accurate picture of the total damage than by simply assuming that each primary knock-on possesses a certain average energy as has usually been done in the past,[2,4] because the nature as well as the amount of damage is found to vary with this energy.

Concerning the damage from a knock-on atom of very high energy, the picture is represented in Fig. 1. Here the metal crystal is represented schematically in two dimensions as a square lattice, the intersections of the background lines representing the lattice sites. Interstitial atoms and vacancies are represented as indicated, as are the paths of the primary knock-on and subsequent secondaries, tertiaries, etc.

In the present paper (1) the persistence of radiation induced interstitials and vacancies and (2) the production of appreciable atomic interchange by thermal spikes are assumed to be two mutually exclusive effects of radiation damage. This is in contradistinction to the usual assumption,[1] that the interstitials and vacancies persist through the duration of the thermal spikes, even though the spikes may disorder an ordered alloy. The assumption that these two effects are mutually exclusive leads to two separate regions along the path of a high-energy primary knock-on, each retaining a different form of damage, as shown in Fig. 1. The high-energy region to the left of point A will retain as interstitial-vacancy pairs all of the displaced atoms produced here, and there will be no appreciable atomic interchange among the remainder of the atoms. On the other hand, the low-energy region to the right will retain essentially none of the interstitial-vacancy pairs produced, but the normal atoms will not retain their respective sites. Thus, if this region were initially part of an ordered superlattice alloy, these atoms will become disordered.

The author was led to make the above assumption by the results of the calculations of the next section. It is found that, for many metals, the thermal spikes reach temperatures well above the normal melting point of the material. Seitz[1] has estimated that the spikes reach

[4] J. D. Ozeroff, KAPL-205, 1949.

temperatures of the order of 10^4 °K for periods of the order of 10^{-11} sec. The present calculations indicate that, in regions which reach temperatures above the normal melting point of the material, the concentration of interstitial atoms and vacant lattice sites is at least several percent of the concentration of normal atoms and lattice sites. Thus, the average separation between interstitials and vacancies in such regions should be only two or three interatomic distances or less. This should give rise to large local strains in the material.

If a true melting point exists for the material which is heated to high temperatures by a thermal spike, it will be at a somewhat higher temperature than the normal melting point at atmospheric pressure, because the material will be held under high pressure by the surrounding undisturbed lattice. However, the temperature of the melting point should not be changed by more than about a factor of 2. Thus temperatures of the order of 10^4 °K are still large compared with the melting point, implying that it is not correct to think of the material as a solid while it is at high temperature. Because of the high pressures and temperatures, it is uncertain whether this material can be more appropriately referred to as a liquid or as a dense gas. In subsequent discussion, the words "melting" and "liquid" will be used to describe this material, but it should be remembered that their meanings are not necessarily conventional.

The large local strains in the material associated with the production of interstitial-vacancy pairs cannot persist indefinitely in the liquid, since the lattice sites surrounding an interstitial atom in a solid no longer exist. Upon melting, an atom, which in the solid was an interstitial, is no longer defined as an interstitial. The region including a few atoms around the location of this atom will contain an abnormally high concentration of atoms, due to the presence of the "extra" atom. Likewise, the region around a vacancy will, after melting, have an abnormally low concentration of atoms. Thus, immediately after melting, the former interstitials and vacancies will be considered to be replaced by local "density fluctuations." If the quenching is extremely rapid, it is expected that these density fluctuations will give rise to an equal number of interstitials and vacancies again upon resolidification. If the liquid state is maintained, these density fluctuations should relax as a result of the associated local strains. If this relaxation time is longer than the time the material remains melted, the radiation-induced, interstitial-vacancy pairs will persist during the melting and resolidification, while if it is much shorter, they will not. The time for relaxation of these strains has been estimated by the author to be of the order of 10^{-12} sec, as the frequency of oscillation of any atom should be of the order of 10^{13} sec^{-1}, and a chain of two or three atoms must participate in the relaxation of a given "density gradient." As this is appreciably shorter than the time the spike remains melted, the interstitials and vacancies are assumed not to persist in a region which has been heated above the melting temperature.

TABLE I.

Z	Element	r_0(A)	E_{tr}(ev)	Z	Element	r_0(A)	E_{tr}(ev)
11	Na	3.708	180	56	Ba	4.34	860
12	Mg	3.190	550	57	Ta	3.73	4200
13	Al	2.856	1200	58	Ce	3.64	5900
19	K	4.618	140	59	Pr	3.633	6000
20	Ca	3.93	420	60	Nd	3.62	6500
21	Sc	3.205	2500	63	Eu	3.960	1500
22	Ti	2.91	5000	64	Gd	3.554	8300
23	V	2.627	9600	65	Tb	3.508	10 000
24	Cr	2.493	15 000	66	Dy	3.499	10 000
26	Fe	2.476	20 000	67	Ho	3.480	11 000
27	Co	2.501	20 000	68	Er	3.459	12 000
28	Ni	2.486	23 000	69	Tm	3.446	13 000
29	Cu	2.551	23 000	70	Yb	3.866	3700
30	Zn	2.659	19 000	71	Tu	3.439	13 000
37	Rb	4.87	150	72	Hf	3.14	33 000
38	Sr	4.30	610	73	Ta	2.854	73 000
39	Y	3.59	3000	74	W	2.734	110 000
40	Zr	3.16	9000	75	Re	2.734	105 000
41	Cb	2.853	25 000	76	Os	2.670	150 000
42	Mo	2.720	36 000	77	Ir	2.709	120 000
44	Ru	2.644	51 000	78	Pt	2.769	110 000
45	Rh	2.685	47 000	79	Au	2.878	80 000
46	Pd	2.745	43 000	81	Te	3.401	16 000
47	Ag	2.882	31 000	82	Pb	3.493	14 000
48	Cd	2.972	21 000	90	Th	3.59	9000

From considerations of ordinary diffusion data, it seems that atomic interchange should not occur appreciably during the short existence of the thermal spike unless the temperature is well above the melting temperature. This leads to the assumption cited earlier regarding two mutually exclusive types of damage. The transition energy possessed by the primary knock-on at point A in Fig. 1, at which the rate of energy loss becomes large enough to anneal the interstitial-vacancy pairs and produce appreciable atomic interchange, has been calculated for most metals, the values being given in Table I.

The region to the right of point A is assumed to have undergone melting and resolidification. It is believed that the atomic interchange occurs during the time the material is melted, predominantly as a result of a certain amount of random motion of the atoms, possibly resembling turbulence in the flow of a liquid, initiated by the relaxation of the local strains when the density fluctuations relax. The stored energy released upon relaxation of these strains will be sufficient to raise the temperature even higher, thereby maintaining the liquid state for a brief period after most of the density fluctuations have disappeared. During this period, the turbulent motion initiated by the relaxation of the strains can presumably continue to a sufficient degree that, upon resolidification, most of the atoms will occupy new lattice sites. Thus, essentially all of the atoms in this region will be "displaced atoms," in the sense that each will be displaced to a new lattice site. A region of crystal which has undergone melting and resolidification in such a manner will therefore be called a "displacement spike."

The resolidification of a displacement spike should occur predominantly on the parent lattice, as it will form an ideal nucleus for crystallization. Thus, the crystal structure of the material should not be destroyed by displacement spikes. A few small microcrystals of entirely new orientation may possibly be formed, however, somewhat as indicated schematically in Fig. 1.

The displacement spikes proposed here should not be confused with the thermal spikes described by Seitz,[1] as they differ in the following aspects. By their definition, displacement spikes cannot sustain radiation-induced, interstitial-vacancy pairs, while it has been assumed that thermal spikes would not anneal the associated interstitial-vacancy pairs. The entire volume of a displacement spike is required to have been melted and resolidified, thus giving it well-defined boundaries, while this is not true for a thermal spike.

Interstitial-vacancy pairs are still produced, according to the present model, in the region to the left of point A in Fig. 1, along with an associated thermal spike. This part of the presently proposed picture is therefore very similar to the model described by Seitz,[1] with the exception that the thermal spike does not reach a temperature at which atomic interchange can occur. The main difference between the model proposed here and the former one, therefore, is the displacement spike concept.

CALCULATIONS

It is desired to calculate the density of interstitial-vacancy pairs produced in the region of the thermal spike along the path of a primary knock-on atom. A "displacement collision" will be defined as a collision of the primary knock-on with a normal lattice atom in which sufficient energy is transferred to the normal atom to separate it from its lattice site, creating an interstitial-vacancy pair. One must then calculate the mean free path of the primary knock-on between successive displacement collisions.

Basic to a calculation of the mean-free path between displacement collisions is a calculation of the cross section for scatter of an atom by an identical atom with an energy transfer greater than a given minimum ϵ_0. This in turn involves a knowledge of the interaction potential energy between two identical atoms $V(r)$. This will be calculated, assuming that the potential of each atom is essentially a screened Coulomb potential of the form

$$\phi(r)=(Zq/r)e^{-r/a}, \tag{1}$$

where q represents the charge of a proton and Z is the atomic number of the atom. From Schiff,[5] one finds that, for moderately heavy atoms, the radius a of the atomic electron cloud is of the order of magnitude

$$a\sim(\hbar^2/mq^2Z^{1/3}),$$

or

$$a=Ca_0Z^{-1/3}, \tag{2}$$

where C is a proportionality factor of the order of unity and

$$a_0=\hbar^2/mq^2=0.5282\times10^{-8}\text{ cm}$$

is the radius of the first Bohr orbit of hydrogen m representing the electronic mass.

Ozeroff[4] has made a calculation of $V(r)$ based on the same model, but his result is apparently in error. The detailed calculations are therefore presented in Appendix A, giving the expression

$$V(r)=(Z^2q^2/r)e^{-r/a}(1-r/2a). \tag{3}$$

This approaches the expected Coulomb repulsion as r approaches 0. At $r=2a$, it changes sign, becoming a weak attractive potential at large distances, with a minimum at $r=a(1+\sqrt{3})$. This is consistent with the model used, in which an atom consists of a nucleus surrounded by the rigid charge distribution ρ, given in Appendix A. It may not be consistent with the true physical picture, however, because effects such as the redistribution of charge during a collision and exchange interactions have been neglected. The closed shell repulsion between ions is probably the largest effect neglected. Certainly, for r somewhat less than $2a$, the above expression is expected to be approximately correct, as the shielded Coulomb repulsion is expected to dominate at short distances.

The potential function, Eq. (1), which has been used, is best suited for heavy un-ionized atoms. For the lighter metals, the quantitative results of such a potential may be inaccurate, because the manner in which the screened Coulomb field drops off does not approximate the exponential form as closely as in the case of the heavier elements. However, the range and strength of the interaction should still be correct as to order of magnitude; and, therefore, it should be possible to draw good qualitative conclusions by use of it, which should be of the correct order of magnitude, even for the very light metals, as long as they are not ionized.

A good approximation to the upper limit of kinetic energy at which a moving atom can be considered un-ionized is

$$E=(M/m)E_{\text{ion}},$$

where M is the mass of the atom, m is the electronic mass, and E_{ion} is the ionization energy of the atom. If it is assumed that an ionization energy of 5 ev is sufficient to ionize any atom to a point such that the above expression for $V(r)$ no longer gives good results, then the following table gives the kinetic energies below which the present treatment is valid.

Atom	Energy, ev
Be^9	80 000
Al^{27}	250 000
Cu^{63}	570 000

In most types of irradiation, the energy transferred to the primary knock-on will be low enough so that the

complete history of the primary knock-on and all successive knock-ons can be treated by use of Eq. (3).

The accuracy of this potential is justified for atomic separations less than about $1.5a$. At distances between about $5a$ and $10a$, the closed shell repulsion is expected to be the dominant interaction. Thus, in this range, the potential is assumed to be an exponential repulsion of the Born-Mayer type,[6] Ae^{-Br}, with constants adjusted to fit observed compressibility data.[7] In the interval $1.5a<r<5a$, the potential will be left unspecified; the assumption will be made, however, that it is a monotonically decreasing function of r in this range connecting smoothly with the assumed expressions at both ends.

Having arrived at a potential function, a satisfactory method for calculating the scattering cross section must now be chosen. The general condition which must be satisfied in order that either the Born approximation or the classical treatment may be valid is

$$\lambda\ll a, \tag{4}$$

where

$$\lambda=h/Mv$$

is the de Broglie wavelength of the moving particle and a represents the order of the dimensions of the scattering field. This condition is fulfilled for atoms with energies greater than 100 ev when the scattering field is given by Eq. (3). Hence, in all collisions involved in the present problem, either the Born approximation or the classical treatment, or a combination of both, can be used according to the conditions described by Williams.[8] The classical treatment can be applied when

$$Vb/\hbar v\gg 1, \tag{5}$$

where b represents the impact parameter in collisions in which the energy transfer is small compared with the energy of the moving particle; V represents the strength of the potential evaluated at this distance, and v is the velocity of the moving particle. In potentials and regions of potentials where the opposite condition is fulfilled, the Born approximation can be applied.

Collisions of interest in the present problem involve energy transfers of between 1 and 25 ev, where the energy of the moving particle is between 1000 and 100 000 ev. Thus, b in Eq. (5) can be considered as truly representing the impact parameter. The relationship between b and ϵ will be determined later, and it will then be shown that Eq. (5) is fulfilled for values of ϵ_0 between 1 and 25 ev as long as $Z>10$. The classical approximation will therefore be used, and it must be remembered that the results will only be good for metals whose atomic number exceeds 10.

In Appendix B, the classical treatment of the problem is carried through, assuming that Eq. (3) represents the correct potential over the entire range of r. On this assumption, the expression for the energy transfer is found to be

$$\epsilon=4\frac{Z^{14/3}}{C^2}\frac{E_R^2}{E}F(b/a), \tag{6}$$

where E_R represents the Rydberg energy 13.52 ev and

$$F(x)\equiv\{K_1(x)-(x/2)K_0(x)\}^2, \tag{7}$$

where K denotes the modified Bessel function of the second kind.

The cross section for scattering with an energy transfer greater than ϵ for such collisions will then be

$$\sigma_\epsilon=\pi b^2=\pi a^2\left\{F^{-1}\left(\frac{C^2E\epsilon}{4E_R^2Z^{14/3}}\right)\right\}^2, \tag{8}$$

where $F^{-1}(x)$ is the inverse of the function defined in Eq. (7).

The function $F(x)$ is shown as the solid line in Fig. 2. According to Eq. (6) this function is proportional to the energy transfer when the impact parameter is ax. It is noticed that at $x=2.4$, the curve drops to 0, so that a particle with an impact parameter equal to $2.4a$ should not be deflected. The reason for this is apparent from the potential used, which has its minimum at $2.73a$. Thus, an impact parameter of $2.4a$ apparently corresponds to the path which a particle must follow if it is acted on equally by the attractive and repulsive forces, giving no net deflection.

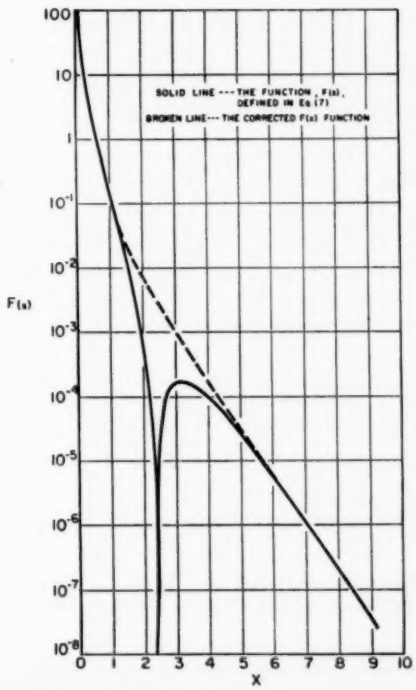


FIG. 2. Function $F(x)$ as defined in Eq. (7) and corrected $F(x)$ function.

[5] L. I. Schiff, *Quantum Mechanics* (McGraw-Hill Book Company, Inc., New York, 1949), p. 168.
[6] M. Born and J. E. Mayer, Z. Physik **75**, 1 (1932).
[7] P. W. Bridgman, *The Physics of High Pressure* (G. Bell and Sons, Ltd., London, 1949), p. 160.
[8] E. J. Williams, Revs. Modern Phys. **17**, 217 (1945).

Consider now the effects of having neglected the closed shell repulsion. This interaction is approximated by a simple exponential form Ae^{-Br}, which is just the negative of the form approached by Eq. (3) for $r \gg 2a$. The energy transfer in such a collision will be independent of the sign of the potential and will depend only on the magnitude. The magnitude of the closed shell repulsion has been calculated for several of the inert gases,[9] and it can be estimated from experimental compressibility data[7] for metals. In general, the evaluation of A and B from compressibility data yields values comparable with those of the attractive term in Eq. (3). Thus, the total interaction at distances greater than about $7a$ is of the same order of magnitude as given by Eq. (3) and is positive rather than negative. The energy transfer for large impact parameter should therefore be of the same magnitude as that calculated here. Thus, for large x, although the sign of the potential is wrong, the $F(x)$ curve in Fig. 2 should be approximately correct.

At distances somewhat less than $2a$, Eq. (3) is expected to represent the true interaction, as the Coulomb interactions should predominate here. The $F(x)$ curve in Fig. 2 should therefore be correct up to about $x=1.5$.

As the true potential is a monotonically decreasing function both at large and small distances, rather than a function whose slope changes sign, as is Eq. (3), there seems to be no further reason for supposing that it has a minimum in the neighborhood of $2a$. It will therefore be assumed that $V(r)$ is a monotonically decreasing function of r throughout the range $0<r<10a$. This will then eliminate the zero in the $F(x)$ curve and demand that $F(x)$ be a monotonically decreasing function of x. It is therefore assumed that the true $F(x)$ curve can be obtained by bridging the gap smoothly, giving the broken line in Fig. 2.

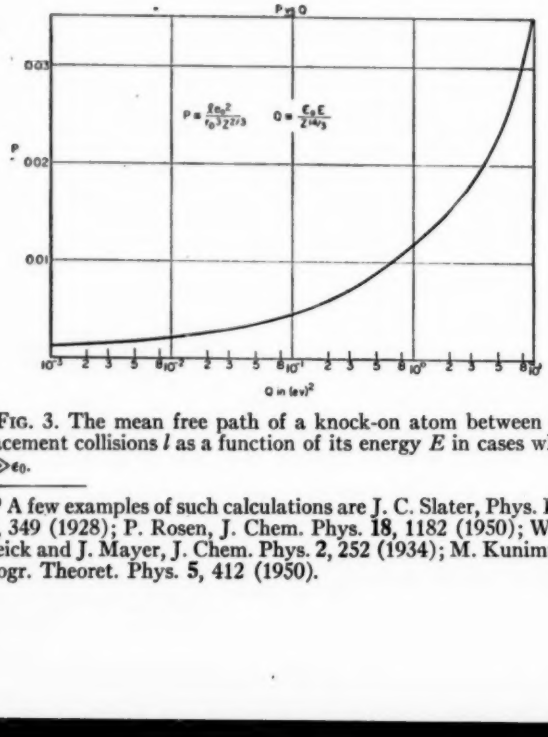


FIG. 3. The mean free path of a knock-on atom between displacement collisions l as a function of its energy E in cases where $E \gg \epsilon_0$.

[9] A few examples of such calculations are J. C. Slater, Phys. Rev. **32**, 349 (1928); P. Rosen, J. Chem. Phys. **18**, 1182 (1950); W. E. Bleick and J. Mayer, J. Chem. Phys. **2**, 252 (1934); M. Kunimune, Progr. Theoret. Phys. **5**, 412 (1950).

It would be possible to estimate an $F(x)$ curve for each metal, basing the right-hand end on the compressibility data for the particular metal. This process is tedious and seems unwarranted as a result of the uncertainty of other approximations which must be made. Rather, the corrected $F(x)$ curve in Fig. 2 will be assumed as an average and will be applied to all metals. Therefore, throughout the remainder of the present paper, and in the use of Eqs. (6) and (8), $F(x)$ will be taken to be defined by the broken curve in Fig. 2, rather than by Eq. (7) which gives the solid curve.

For purposes of determining whether or not Eq. (5) is fulfilled, the form of the assumed potential at large distances can be taken approximately as the magnitude of the attractive part of Eq. (3),

$$V(r) \rightarrow \frac{Z^2q^2}{2a} e^{-r/a}. \tag{9}$$

Using this and Eq. (6), it can now be seen that Eq. (5) is fulfilled if $\epsilon_0 E > 1000$ (ev)^2 and $Z > 10$.

The mean free path between displacement collisions is given by

$$l = 1/N_0 \sigma_{\epsilon_0}, \tag{10}$$

where σ_{ϵ_0} is the cross section for a displacement collision and N_0 is the density of atoms. It will be assumed that, for both close-packed and body-centered cubic metals, N_0 is given by

$$N_0 = 1.4/r_0^3, \tag{11}$$

where r_0 is the interatomic distance. This is accurate only for close-packed metals, but is less than 8 percent in error for the body-centered cubic structure. Equation (8) will be used for σ_{ϵ_0}. One can then pick any metal with a given Z and r_0, assume a value for the displacement energy ϵ_0 and plot l *vs* E.

A more general curve can be obtained, however, for all metals by plotting the quantities P and Q, which are defined by the expressions

$$P \equiv \frac{la_0^2}{r_0^3 Z^{2/3}}, \tag{12}$$

$$Q \equiv \frac{E\epsilon_0}{Z^{14/3}}. \tag{13}$$

Then, from Eq. (8) it is seen that if one sets

$$F(x) = (C^2/4E_R^2)Q, \tag{14}$$

the cross section for displacement is

$$\sigma_{\epsilon_0} = \pi a^2 x^2 = \frac{\pi C^2 a_0^2}{Z^{2/3}} x^2. \tag{15}$$

Therefore, combining Eqs. (10), (11), (12), and (15),

$$P = 1/1.4\pi C^2 x^2. \tag{16}$$

The value of C will be taken as 2.09, as used by Ozeroff[4] to agree with the Thomas-Fermi atom model. Then,

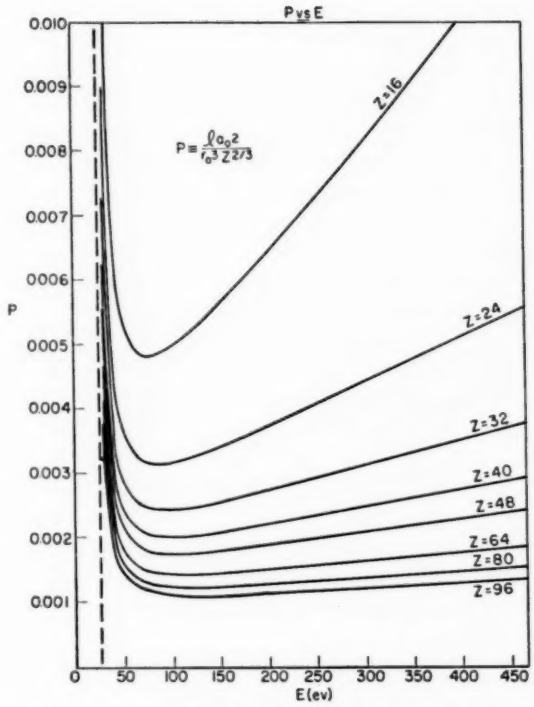


FIG. 4. The mean free path of a knock-on atom between displacement collisions l as a function of its energy E in cases where $E \sim \epsilon_0$.

from Eqs. (14) and (16) and the $F(x)$ curve, one can plot P *vs* Q, giving the curve in Fig. 3. For any given metal, one knows Z and r_0; and, by estimating a value for ϵ_0, one obtains the l *vs* E curve for a particular metal from Eqs. (12) and (13) as P and Q are directly proportional to l and E, respectively.

The impulse approximation was used to obtain the P *vs* Q plot in Fig. 3. It is therefore possible to set a lower limit on the values of Q for which this curve is valid, by applying the validity criterion for the impulse approximation, namely,

$$\epsilon_0 \ll E.$$

This gives

$$Q \gg \epsilon_0^2/Z^{14/3}. \tag{17}$$

To obtain a rough idea of the form of the l *vs* E curve for lower values of E, it will be assumed that the scattering can be treated as hard sphere scattering, the cross section for scattering with an energy transfer greater than ϵ_0 being given approximately by

$$\sigma\epsilon_0 = \pi b_0^2[1-(\epsilon_0/E)], \tag{18}$$

where b_0 represents the distance of closest approach of the atoms in a head-on collision.

To obtain an expression for b_0, the form of the potential at large distances will be taken as that given in Eq. (9). Setting this equal to E, one can solve for b_0, giving

$$b_0 = \frac{Ca_0}{Z^{1/3}} \ln\left(\frac{Z^{7/3}}{C}\frac{E_R}{E}\right). \tag{19}$$

Combining Eqs. (10), (11), (12), (18), and (19), it is found that in this low-energy limit, P is given by

$$P = \frac{1}{1.4\pi C^2\left(1-\frac{\epsilon_0}{E}\right)\left[\ln\left(\frac{Z^{7/3}}{C}\frac{E_R}{E}\right)\right]^2}. \tag{20}$$

In this case, it is not possible to choose a single coordinate proportional to E which will give a single curve representing all metals. Thus, a separate curve must be plotted for each set of values for Z and ϵ_0. Figure 4 has been plotted showing the family of l *vs* E curves which one obtains by choosing ϵ_0 as 25 ev. The curves from Figs. 3 and 4 have been combined to give the l *vs* E curve for copper as a typical example in Fig. 5. While the two curves do not join exactly, the agreement between them is satisfactory for present purposes.

The displacement energy ϵ_0 assumed to be 25 ev, is in general of the order of, or less than, 100 times the energy per atom necessary to melt the material within a displacement spike. From this fact, one can see from Eq. (8) and Fig. 2 that the number of displaced atoms in any region should be about one-fourth the number of atoms which are heated to the melting temperature. Thus, in regions in which the lattice is heated to sufficiently high temperatures that melting may occur, the density of displaced atoms according to this theory should be about 25 percent of the total atom density. While this figure may be somewhat in error as a result of the rather crude approximations made in the preceding calculations, it seems certain that the number should be at least 5 percent, supporting the assumption that radiation induced Frenkel defects cannot persist in a region which has been melted and resolidified.

A criterion must now be set up to determine whether or not such melted regions are produced, and what the transition energy E_{tr} of the knock-on is when such production starts. There are three factors which contribute to the heating of a displacement spike: (1) the nondisplacement elastic collisions of the knock-on atoms with the normal atoms of the lattice; (2) the transfer of

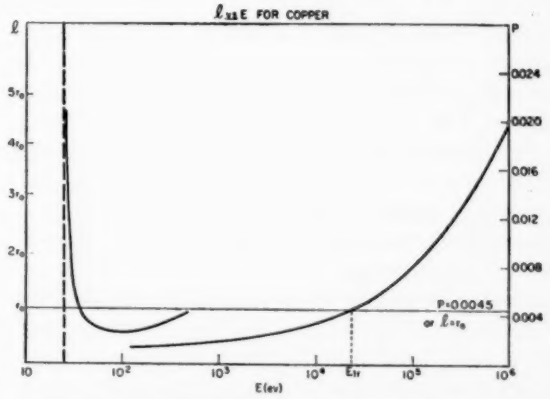


FIG. 5. The mean free path of a copper knock-on atom between displacement collisions l as a function of its energy E.

energy to the lattice from the electrons which have been excited by the knock-on atoms; and (3) the release of energy as a result of the annealing of the Frenkel defects. Calculations made by Brooks[10] show that the energy imparted to the electrons by a knock-on atom is dispersed among a large number of electrons before it can produce an appreciable temperature rise in the atomic lattice; hence Process (2) is negligible. Process (3) occurs as a result of Process (1); therefore, the heating as a result of Process (1) alone must be great enough to initiate the action which is then continued by Process (3). It will therefore be assumed that the lattice heating resulting only from nondisplacement elastic collisions of the primary knock-on with lattice atoms must be sufficient to melt a continuous cylindrical region along its path, from two to four interatomic distances in diameter, in order to initiate the production of a displacement spike. Thus, the value of E_{tr} will be taken as the energy of the primary knock-on at which the rate of energy loss to nondisplacement elastic collisions is just large enough to heat to melting temperature from 4 to 12 lattice atoms per interatomic distance along its path.

This criterion can be simplified somewhat by considering the previously estimated ratio of the relative numbers of displaced and melted atoms. One can therefore estimate that the value of E_{tr} is the energy of the primary knock-on at which the mean free path between successive displacement collisions becomes of the order of one interatomic distance. This will be taken as the criterion for the determination of E_{tr}. It is therefore seen that E_{tr} for copper will be given by the second intersection of the curve in Fig. 5 with the horizontal broken line, which represents a mean-free path between displacement collisions equal to one interatomic distance.

Similarly, values of E_{tr} for all metals can be obtained from Fig. 3 by calculating the value of P corresponding to one interatomic distance in the given metal, and determining E_{tr} from the corresponding value of Q, given by the curve. Table I has been calculated by this method, assuming ϵ_0 to be 25 ev for each metal.

THE NATURE OF DISPLACEMENT SPIKES

From Table I, one deduces that, if the radiation damage model proposed in this paper is correct, the damage produced in the heavier metals by most of the common kinds of irradiation should consist primarily of displacement spikes. Only in a few of the lightest metals should the production of interstitial-vacancy pairs be appreciable. One likely means of experimentally differentiating between this model and the former one should be a determination of the number of interstitial-vacancy pairs produced in some of the heavier metals by pile neutron irradiation. The present model predicts few or none, while other workers have assumed this to be the primary damage resulting from such irradiation. Another difference is that the former model assumes that the damage produced in both the light and heavy elements is of essentially the same nature, the primary effect being the production of interstitial-vacancy pairs. The presently proposed model, however, predicts that the damage in the heavy elements will be of a different nature than in the light elements, consisting chiefly of displacement spikes. In this connection, some speculation on the nature of displacement spikes and on the differences expected to be observed in the physical effects resulting from displacement spikes and from interstitial-vacancy pairs seems appropriate.

It is easy to imagine processes by which dislocation loops can be formed within a displacement spike during its resolidification. The formation of small microcrystals of new orientation, as illustrated in Fig. 1, should be a more difficult process and, if it occurs, should be much less frequent than the production of dislocation loops. The reason for this is that the boundary of such a small microcrystals should consist of an array of these dislocation loops. Thus it seems that small dislocation loops should be one of the primary products of displacement spikes.

As a result of the small size of these dislocation loops, they will be in strong tension; and, as a result, they should anneal easily, simply by collapsing. The activation energy for annealing should increase with the size of the loop. Thus, the annealing of this type of damage can be expected to have a variable activation energy, giving an annealing process which occurs over a rather wide range of temperature. It should be possible to observe annealing beginning at very low temperatures, but which will not run to completion until the temperature is raised considerably higher.

This is in contrast to the annealing which one should observe for a process involving only a single activation energy. Here, if a temperature is found at which annealing will begin, the process should run to completion at this temperature. From measurements of the annealing of radiation damage, it may therefore be possible to separate the two models, because the annealing of interstitials and vacancies should be characterized by a small number of discrete activation energies rather than a continuous range.

The size of displacement spikes can probably be estimated crudely by dividing the energy available for production of the spike, which is just the energy of the primary knock-on at the time spike production starts, by an average energy per atom when the spike is in the melted condition. The energy per atom necessary to melt typical metals at atmospheric pressure is between 0.1 ev and 0.2 ev. The metal within displacement spikes will be held at high pressure by the surrounding lattice, raising the melting point somewhat, to a value which should still be less than 0.5 ev. To account in a rough manner for heat losses to the surrounding nonmelted lattice and to the electronic system, the estimate will arbitrarily be raised to the order of 1.0 ev per atom as the amount of available energy dis-

[10] H. Brooks (private communication).

sipated by each atom of the spike. Using this figure, one obtains an average size for the displacement spikes produced by 2-Mev neutrons in copper of about 2×10^4 atoms. If the region is assumed spherical, this represents a sphere about 75A in diameter. This then also represents an approximate upper limit on the size of dislocation loops which can be produced.

It was argued earlier that a density of interstitial-vacancy pairs giving an average vacancy-interstitial separation of only two or three interatomic distances would not persist throughout the associated thermal spike, because the relaxation time of the pairs was probably short relative to the melted period of the spike. It is possible, however, that a lower concentration of interstitial-vacancy pairs could persist during the temperature pulse. In particular, one might think that an average vacancy-interstitial separation of 8 to 10 interatomic distances might have a relaxation time somewhat longer than the time the spike remains melted. Relaxation of the interstitial-vacancy pairs therefore may be thought of as progressing until the concentration of atoms in interstitial positions is decreased to the order of 10^{-3}, at which time resolidification occurs, freezing in the remaining interstitial-vacancy pairs. Hence, it is possible that a few interstitial-vacancy pairs will be produced in each displacement spike, the maximum concentration being estimated to be of the order of 10^{-3}. For pile neutron irradiation on copper, this corresponds to about 20 interstitial-vacancy pairs per primary interaction, a considerably smaller number than one would obtain from calculations based on the earlier model, of the type made by Seitz.[2] In the lighter elements, Be, C, Na, Mg, and Al, the two models should predict about the same type of damage, as the size of the displacement spikes in such cases should be negligible. Attempts to experimentally differentiate between the two models should therefore be carried out on the heavier metals.

CONCLUSIONS

It would be desirable to measure a direct effect of irradiation on some physical property of the heavier metals which could be definitely assigned to either interstitials and vacancies or to dislocation loops. Such an effect, however, may be difficult to find, as both should increase the electrical resistivity, both are expected to produce increases in hardness, and other effects are probably common to both. Two more indirect methods have been described in the last section by which the accuracy of the present model may be checked. These are (1) an analysis of the dependence of the nature of the damage on atomic number and (2) an analysis of the annealing of property changes produced by the damage. It is suggested that such experiments be carried out on metals irradiated with pile neutrons at temperatures as low as possible. The average size of displacement spikes produced by pile neutrons should be considerably larger than those produced by cyclotron irradiation. Thus the production of dislocation loops and other displacement spike effects should be more pronounced in neutron irradiation. In order to retain as much of the damage as possible, the metals should be held at temperatures as low as can be maintained during irradiation, because, according to the present model, displacement spike effects may begin to anneal at temperatures below that of liquid nitrogen.

ACKNOWLEDGMENTS

The author wishes to express his gratitude to the following persons for many valuable suggestions which have been incorporated in this paper: D. B. Bowen, C. E. Dixon, J. S. Lomont, M. M. Mills, and F. Seitz.

APPENDIX A

Consider a system of two atoms. The calculation is simplified considerably by imagining the atoms to be nonidentical at first. The separation distance will be denoted by $\mathcal{R}$, the first being located at coordinates $\mathbf{R}_1$ and the second at $\mathbf{R}_2$. The respective potentials will be

$$\phi_1=Z_1q\frac{\exp(-|\mathbf{r}-\mathbf{R}_1|/a_1)}{|\mathbf{r}-\mathbf{R}_1|},$$

$$\phi_1=Z_2q\frac{\exp(-|\mathbf{r}-\mathbf{R}_2|/a_2)}{|\mathbf{r}-\mathbf{R}_2|},$$

and as

$$\rho=-\frac{1}{4\pi}\nabla^2\phi,$$

the associated charge distributions are

$$\rho_1=-\frac{Z_1q}{4\pi}\left\{\frac{\exp(-|\mathbf{r}-\mathbf{R}_1|/a_1)}{a_1^2|\mathbf{r}-\mathbf{R}_1|}-4\pi\delta^{(3)}(\mathbf{r}-\mathbf{R}_1)\right\},$$

$$\rho_2=-\frac{Z_2q}{4\pi}\left\{\frac{\exp(-|\mathbf{r}-\mathbf{R}_2|/a_2)}{a_2^2|\mathbf{r}-\mathbf{R}_2|}-4\pi\delta^{(3)}(\mathbf{r}-\mathbf{R}_2)\right\},$$

where $\delta^{(3)}$ is the three-dimensional δ function representing the nuclear point charge. The electrostatic energy of the system is given by

$$W=\frac{1}{2}\int_{\text{all space}}(\phi_1+\phi_2)(\rho_1+\rho_2)d^{(3)}\mathbf{r}.$$

The cross terms will give the interaction energy

$$V=\frac{1}{2}\int_{\text{all space}}(\phi_1\rho_2+\phi_2\rho_1)d^{(3)}\mathbf{r}.$$

The two terms in this integral must be equal as a consequence of Green's reciprocation theorem,[11] and therefore

$$V=\int_{\text{all space}}\phi_1\rho_2d^{(3)}\mathbf{r}.$$

[11] W. R. Smythe, *Static and Dynamic Electricity* (McGraw-Hill Book Company, Inc., New York, 1950), p. 34.

An equivalent result should be obtained if $\phi_2\rho_1$ is integrated instead of $\phi_1\rho_2$. Thus,

$$\frac{V}{Z_1Z_2q^2}=\frac{\exp(-\Re/a_1)}{\Re}-\frac{1}{4\pi a_2^2}\times\int_{\text{all space}}\frac{\exp(-|\mathbf{r}-\mathbf{R}_1|/a_1-|\mathbf{r}-\mathbf{R}_2|/a_2)}{|\mathbf{r}-\mathbf{R}_1||\mathbf{r}-\mathbf{R}_2|}d^{(3)}r$$
$$=\frac{\exp(-\Re/a_2)}{\Re}-\frac{1}{4\pi a_1^2}\times\int_{\text{all space}}\frac{\exp(-|\mathbf{r}-\mathbf{R}_1|/a_1-|\mathbf{r}-\mathbf{R}_2|/a_2)}{|\mathbf{r}-\mathbf{R}_1||\mathbf{r}-\mathbf{R}_2|}d^{(3)}r,$$

and therefore

$$\int_{\text{all space}}\frac{\exp(-|\mathbf{r}-\mathbf{R}_1|/a_1-|\mathbf{r}-\mathbf{R}_2|/a_2)}{|\mathbf{r}-\mathbf{R}_1||\mathbf{r}-\mathbf{R}_2|}d^{(3)}r=4\pi a_1^2a_2^2\frac{\exp(-\Re/a_1)-\exp(-\Re/a_2)}{\Re(a_1^2-a_2^2)}.$$

Thus,

$$V(\Re)=\frac{Z_1Z_2q^2}{\Re}\left\{\frac{a_1^2\exp(-\Re/a_2)-a_2^2\exp(-\Re/a_1)}{\Re a_1^2-a_2^2}\right\}$$

for $a_1\neq a_2$.

The atoms can now be considered identical again, setting

$$a_1=a_2=a,\quad Z_1=Z_2=Z.$$

$V(\Re)$ then takes on an indeterminate form, which when evaluated becomes

$$V(\Re)=\frac{Z^2q^2}{\Re}\exp(-\Re/a)\left(1-\frac{\Re}{2a}\right).$$

APPENDIX B

As ϵ_0 is generally small relative to the energy of the moving particle, the classical treatment can be simplified by use of the "impulse approximation," in which the moving atom is considered undisturbed by the stationary atom. The impulse given to the stationary particle is given by

$$I=\int_{-\infty}^{+\infty}F_\perp dt=2\int_0^{+\infty}\left(\frac{M}{2E}\right)^{\frac{1}{2}}Fedx,$$

where $F_\perp$ is the component of the force exerted on the stationary atom by the moving atom perpendicular to the path of the moving atom, and is given by

$$F_\perp=F_r\sin\theta,\quad F_r=-\partial V/\partial r.$$

Here, θ is the angle between the line joining the two atoms and the path of the moving atom, and V is given by Eq. (3). This gives

$$\frac{I}{Z^2q^2}=b\left(\frac{2M}{E}\right)^{\frac{1}{2}}\{\vartheta_1+\vartheta_2+\vartheta_3\},$$

where

$$\vartheta_1\equiv-\frac{1}{2a^2}\int_b^\infty\frac{e^{-r/a}}{(r^2-b^2)^{\frac{1}{2}}}dr,$$

$$\vartheta_2\equiv-\frac{1}{a}\int_b^\infty\frac{1}{r}\frac{e^{-r/a}}{(r^2-b^2)^{\frac{1}{2}}}dr,$$

$$\vartheta_3\equiv\int_b^\infty\frac{1}{r^2}\frac{e^{-r/a}}{(r^2-b^2)^{\frac{1}{2}}}dr,$$

and b is the impact parameter.

$-2a^2\vartheta_1$ is just the Laplace transform of the function

$$\begin{cases}0 & \text{for } 0<r<b\\ \dfrac{1}{(r^2-b^2)^{\frac{1}{2}}} & \text{for } b<r<\infty,\end{cases}$$

which is equal to $K_0(b/a)$, where K denotes the modified Bessel function of the second kind. The integrals ϑ_2 and ϑ_3 can then be obtained by integrating ϑ_1 with respect to $1/a$. Thus,

$$\vartheta_1=-\frac{1}{2a^2}K_0(b/a),$$

$$\vartheta_2=-\frac{1}{a}\int_{1/a}^\infty K_0(xb)dx,$$

$$\vartheta_3=\int_{1/a}^\infty\int_y^\infty K_0(xb)dxdy.$$

By use of the formula

$$\int_a^\infty dy\int_y^\infty f(x)dx=-\int_a^\infty(a-x)f(x)dx.$$

ϑ_3 can be shown to be equal to

$$\vartheta_3=-\int_{1/a}^\infty[(1/a)-x]K_0(xb)dx.$$

Thus

$$\vartheta_2+\vartheta_3=\int_{1/a}^\infty xK_0(xb)dx=(1/ab)K_1(b/a),$$

and

$$I=\frac{Z^2q^2}{a}\left(\frac{2M}{E}\right)^{\frac{1}{2}}\{K_1(b/a)-\tfrac{1}{2}(b/a)K_0(b/a)\}.$$

The energy transfer is

$$\epsilon=\frac{I^2}{2M}=4\frac{Z^{14/3}}{C^2}\frac{E_R^2}{E}F(b/a),$$

where C is defined in Eq. (2), and

$$E_R\equiv q^2/2a_0=13.52\text{ ev}=\text{Rydberg energy},$$

and

$$F(x)\equiv\{K_1(x)-(x/2)K_0(x)\}^2.$$

A Capillary Viscometer with Continuously Varying Pressure Head

SAMUEL H. MARON, IRVIN M. KRIEGER, AND ARTHUR W. SISKO
Department of Chemistry and Chemical Engineering, Case Institute of Technology, Cleveland, Ohio
(Received January 7, 1954)

A new capillary viscometer has been designed to study the flow behavior of non-Newtonian fluids. The instrument is convenient, absolute, and accurate, and covers continuously a wide shearing stress range in a single determination. A falling mercury column forces the sample through the capillary. Measurement of the column height as a function of time gives both pressure drop and flow rate. Provision is also made for the use of driving fluids less dense than mercury for measurements at lower shearing stresses. The density of the sample need not be determined. Methods and illustrations are given of the application of the new instrument to the determination of the absolute viscosities of Newtonian liquids, and the flow curves of non-Newtonian fluids.

INTRODUCTION

CAPILLARY viscometers find very wide application for the measurement of the flow characteristics of fluids. In the case of Newtonian fluids the observation of the rate of flow at a single pressure drop across the capillary is sufficient to define the flow behavior. However, in the case of non-Newtonian fluids a range of shearing stresses must be employed, and this is usually obtained by varying the pressure drop. An effective means of accomplishing this pressure variation was devised by Ostwald and Auerbach,[1] who allowed the liquid to drain from a buret under its own pressure through a capillary, while the times corresponding to various buret readings were recorded. From these data could be calculated the pressure drops and the flow rates. By this means pressure heads ranging from the height of the buret down to zero could be obtained in a single determination.

Giesy and Arzoomanian,[2] working with stiff pastes, extended the pressure range and eliminated drainage errors involved in the Ostwald-Auerbach viscometer by using mercury as a drive fluid to force the sample from a reservoir through a horizontal capillary.

The viscometer described in this paper is a modification of the previously mentioned buret-capillary instruments in the direction of increased precision, versatility, and convenience of operation. Also presented below is a novel method of treating the data for recovery of the rate of shear-shearing stress relation of the fluid studied.

DESCRIPTION AND OPERATION OF NEW VISCOMETER

A sketch of the new viscometer is shown in Fig. 1. The apparatus, constructed entirely of glass, consists of a capillary unit I and a manometer unit II connected by a ball and socket joint *E*. The capillary unit is composed of a water jacket *F* inside of which are sealed the bulb *C* and tube *D* of the same diameter, joined by the capillary *A* and sidearm tube *B*. The function of the sidearm is to facilitate filling of the unit with sample, and to speed attainment of hydrostatic balance by bypassing the capillary. During a run the sidearm stopcock is kept closed.

The manometer unit II consists of two precision-bore tubes, *ca* 0.34 cm in diameter and 100 cm long, a bulb *W*, and a 20-ml hypodermic syringe joined to the manifold as indicated. Tube *m* is used for mercury, while tube *w* is for less dense driving fluids. The manometer tubes are placed against a scale of millimeter graph paper, protected with cellulose acetate film, to permit reading of the column heights. The syringe is used to adjust the amount of mercury in the system and the mercury levels. All the stopcocks shown are large bore and fitted with pressure adapters. As may be seen from

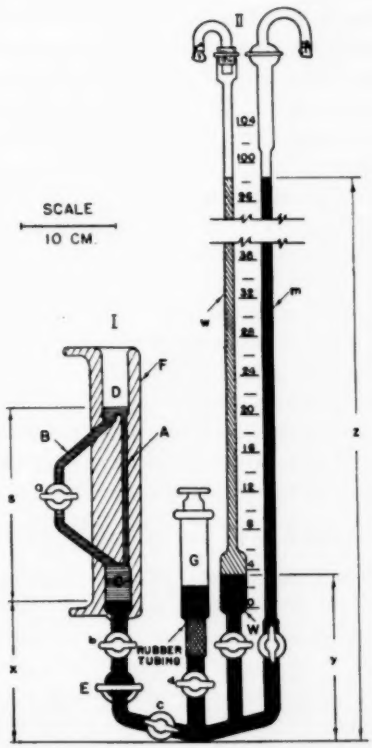


FIG. 1. Diagram of new viscometer.

[1] W. Ostwald and R. Auerbach, Kolloid Z. 41, 56 (1927).
[2] P. M. Giesy and S. Arzoomanian, *Colloid Symposium Monograph* (Chemical Catalog Company, Inc., New York, 1928), Vol. V, p. 253.

the drawing, only mercury comes in contact with the sample even when a different driving fluid is employed.

To permit variation of capillary size, four units such as I were constructed. These differed only in the bore of the capillaries sealed in, and they could readily be interchanged through the ball and socket connection *E*.

To operate the instrument, the units are assembled and held in place by means of a clamp applied at the ball joint, and the stopcocks and syringe are manipulated to raise the mercury in unit I to the top of bulb *C*. Approximately 20 ml of sample are then introduced through the top of unit I, and drawn into bulb *C* and the sidearm by retraction of the syringe. The syringe is then advanced to force out any air present and to fill the capillary with sample, the levels of mercury and sample are adjusted to a point such as shown in Fig. 1, and stopcocks *a* and *b* are closed. Finally, pressure is applied to the syringe to fill the appropriate manometer column with fluid, and stopcock *d* is closed. At this stage of the operation only stopcock *b* prevents the driving fluid from pushing the sample through the capillary.

After the sample has been brought to equilibrium with the water circulating through *F* from a thermostat, stopcock *b* is opened, and the times *t* are noted at which the drive fluid level passes various heights *z* in the manometer column. For this purpose a stopwatch with a "split second" following hand greatly facilitates the timing. When the desired range of pressures has been covered, stopcock *a* is opened, and the height z_0 to which the drive fluid falls at hydrostatic balance is read. An accurate determination of z_0 is necessary since this figure enters into the calculation of the effective driving height. With stopcock *a* closed, z_0 is approached asymptotically with time, and a long period elapses before attainment of hydrostatic balance. With the aid of the sidearm, however, this balance is reached rapidly, and z_0 can be read without waiting.

The sample may at this stage be returned to the lower bulb, and the determination repeated as many times as desired.

At the end of an experiment the syringe is used to draw mercury from the sample bulb to just above stopcock *b*, stopcocks *b* and *c* are closed, and the capillary unit removed for cleaning. By this procedure only 2–3 ml of mercury are contaminated per determination. The capillary units are readily cleaned with either hot nitric acid or dichromate-sulfuric acid cleaning solution.

PRINCIPLE OF OPERATION

Consider first the operation of the instrument with mercury as the drive fluid. Let r_m be the radius of manometer tube *m*, and let *A* be a capillary of radius *R* and length *L*. Also, let the radius of bulbs *C* and *D*, which are of the same diameter, be R_C. Since bulbs *C* and *D* have the same diameter, the height *s* of the sample column, Fig. 1, remains constant during a determination.

In Fig. 1 let *x* and *z* be the heights of the mercury columns in the *U*-tube system above a horizontal base line as indicated. When the sample is flowing through the capillary tube with a volume rate of flow *Q*, the pressure drop *P* across the capillary is

$$P=z\rho_m g-(x\rho_m g+s\rho_s g) \tag{1}$$

where ρ_m and ρ_s are, respectively, the densities of mercury and sample, while *g* is the acceleration due to gravity. When the system is in hydrostatic balance $x=x_0$ and $z=z_0$, and thus

$$O=z_0\rho_m g-(x_0\rho_m g+s\rho_s g). \tag{2}$$

Subtracting Eq. (2) from Eq. (1), we get

$$P=[(z-z_0)-(x-x_0)]\rho_m g. \tag{3}$$

Since the volume of mercury in the system remains constant,

$$x-x_0=-\frac{r_m^2}{R_C^2}(z-z_0) \tag{4}$$

and hence Eq. (3) becomes

$$P=(z-z_0)\rho_m g\left(1+\frac{r_m^2}{R_C^2}\right). \tag{5}$$

Letting now $h=z-z_0$,

$$P=h\rho_m g\left(1+\frac{r_m^2}{R_C^2}\right) \tag{6}$$

or

$$P=\alpha_m h \tag{7}$$

where

$$\alpha_m=\rho_m g\left(1+\frac{r_m^2}{R_C^2}\right). \tag{8}$$

The volume rate of flow is given by

$$Q=-\pi r_m^2\frac{dh}{dt}. \tag{9}$$

From Eqs. (6) and (9) it is evident that determination of *h* as a function of *t*, as the sample flows through the capillary under the driving pressure of the manometer, allows calculation of both *P* and *Q*. It should also be observed that the density of the sample ρ_s does not enter the equations, and hence need not be known.

When an alternate driving fluid of density ρ_w is used in manometer *w* of radius r_w, the pressure drop is

$$P=(z-y)\rho_w g+y\rho_m g-s\rho_s g-x\rho_m g. \tag{10}$$

At hydrostatic balance, $P=0$, and hence

$$O=(z_0-y_0)\rho_w g+y_0\rho_m g-s\rho_s g-x_0\rho_m g. \tag{11}$$

Where y_0 is the value of *y*, the height of mercury in bulb *W*, at hydrostatic equilibrium. Subtracting Eq. (11) from Eq. (10), we obtain

$$P=(z-z_0)\rho_w g+(y-y_0)(\rho_m-\rho_w)g-(x-x_0)\rho_m g. \tag{12}$$

Setting again $z-z_0=h$, and noting that

$$y-y_0=(z-z_0)\frac{r_w^2}{R_W^2}=h\frac{r_w^2}{R_W^2}, \tag{13}$$

where R_W is the radius of the bulb W. Thus,

$$-(x-x_0)=(z-z_0)\frac{r_w^2}{R_C^2}=h\frac{r_w^2}{R_C^2}, \tag{14}$$

Eq. (12) leads to

$$P=\rho_w gh\left\{1+\frac{r_w^2}{R_W^2}\left[\frac{\rho_m}{\rho_w}\left(1+\frac{R_W^2}{R_C^2}\right)-1\right]\right\}, \tag{15}$$

or

$$P=\alpha_w h, \tag{16}$$

where

$$\alpha_w=\rho_w g\left\{1+\frac{r_w^2}{R_W^2}\left[\frac{\rho_m}{\rho_w}\left(1+\frac{R_W^2}{R_C^2}\right)-1\right]\right\}. \tag{17}$$

When $R_W=R_C$, this equation simplifies to

$$\alpha_w=\rho_w g\left[1+\frac{r_w^2}{R_W^2}\left(\frac{2\rho_m}{\rho_w}-1\right)\right], \tag{18}$$

and this is the value of α_w to be used in Eq. (16). As before the rate of flow is given by

$$Q=-\pi r_w^2\frac{dh}{dt}. \tag{19}$$

From Eqs. (16) and (19) it may be seen that again Q and P follow from measurements of h and t, and that again the density of the sample is not needed.

FLOW OF NEWTONIAN FLUIDS IN VISCOMETER

According to Poiseuille's equation the viscosity η of a Newtonian fluid is given by

$$\eta=\frac{\pi R^4 P}{8LQ}. \tag{20}$$

Substituting the relations for P and Q deduced above, we get

$$\eta=-\frac{R^4\alpha}{8Lr^2}\cdot\frac{1}{\dfrac{d\ln h}{dt}} \tag{21}$$

or

$$\frac{d\log_{10}h}{dt}=m=-\frac{R^4\alpha}{8(2.303)Lr^2}\cdot\frac{1}{\eta}=-\frac{B}{\eta} \tag{22}$$

where, depending on the drive system used, $\alpha=\alpha_m$ or α_w and $r=r_m$ or r_w. From Eq. (22) it follows that for a Newtonian fluid a plot of $\log_{10}h$ *vs* t should be linear,

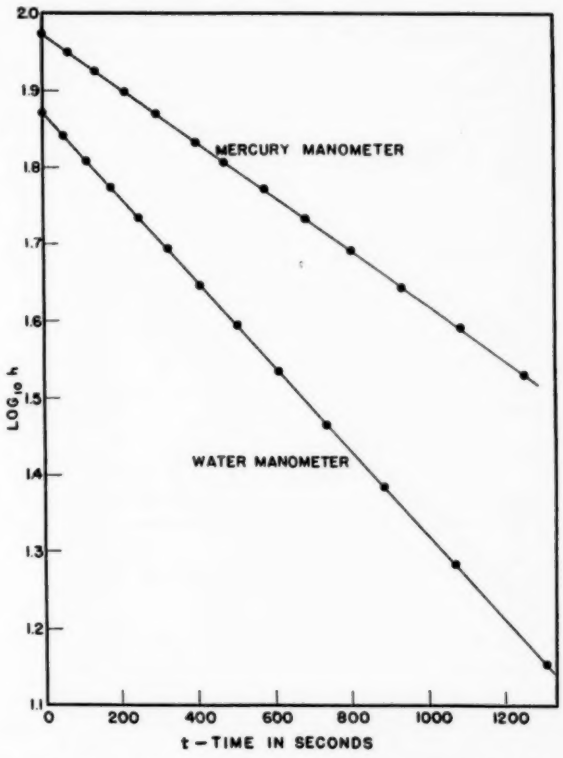


FIG. 2. Typical plots of $\log_{10}h$ *vs* time for Newtonian liquids (white oils).

and the viscosity should be obtainable from the slope of this line and the instrumental parameters.

Figure 2 shows two typical plots of $\log_{10}h$ *vs* t obtained with Newtonian fluids. Table I, in turn, lists data observed on the flow of a number of liquids of known viscosity in four different capillaries with both mercury and water manometers. In all instances slopes were determined by least squares. According to Eq. (22) the product $-\eta m$ should be a constant B for each capillary, and this is actually the case.

The constants B can also be obtained from the instrumental dimensions. For this purpose capillary radii were determined by partially filling the bore with mercury, measuring the length of the mercury thread with a cathetometer, and then weighing the mercury. The radii of the manometers and bulbs were found by introducing weighed quantities of mercury into these, and measuring the differences in levels produced. Finally, the lengths of the capillaries were obtained by filling them with mercury, and measuring the electrical resistance with a Mueller bridge. In these operations corrections were made for the resistances of the leads.

The constants B deduced from instrumental dimensions are given in the sixth and ninth columns of Table I. The agreement between these and $-\eta m$ is excellent, being in all instances but one better than 0.5 percent. Consequently, within the indicated accuracy the instrument may be operated as an absolute viscometer.

TABLE I. Flow data for Newtonian liquids at 30°C.

Capillary	R cm $\times 10^5$	L cm	η poises	Mercury drive B exp $\times 10^5$ $(-\eta m)$	Mercury drive B calc $\times 10^5$	Mercury drive % Deviation	Water drive B exp $\times 10^5$ $(-\eta m)$	Water drive B calc $\times 10^5$	Water drive % Deviation
			0.007982	4.096[a]			4.371	4.354	+0.39
A	1.233	14.59	0.02607	4.090					
			0.1164	4.090					
				av 4.092	4.112	−0.49			
			0.1164	26.02			27.45	27.52	−0.25
B	1.948	14.47	0.2145	25.91					
			0.4213	25.96					
				av 25.96	25.98	−0.08			
			0.7249	157.2			167.3	166.7	+0.36
C	3.033	14.14	0.9269	157.4					
				av 157.3	157.8	−0.32			
			0.9269	394.5					
D	3.818	14.09	1.0086	394.2			419.4	416.6	+0.67
				av 394.4	394.1	+0.08			

[a] Kinetic energy correction applied. For application of this correction see Appendix.

FLOW OF NON-NEWTONIAN FLUIDS IN VISCOMETER

In Fig. 3 is shown a plot of $\log_{10}h$ *vs* t obtained with a 62.2 percent solids synthetic latex at 30°C. The non-linearity of the graph indicates that the latex is non-Newtonian. The problem involved, then, is the computation of the rate of shear-shearing stress relation of the fluid from these data. Two procedures may be employed for this purpose. The first, or incremental one,[3] involves definition of a mean rate of flow $\bar{Q}$ as

$$\bar{Q}=-\pi r^2\frac{h_2-h_1}{t_2-t_1} \tag{23}$$

and a logarithmic mean pressure head

$$\bar{h}=\frac{h_2-h_1}{\ln\dfrac{h_2}{h_1}} \tag{24}$$

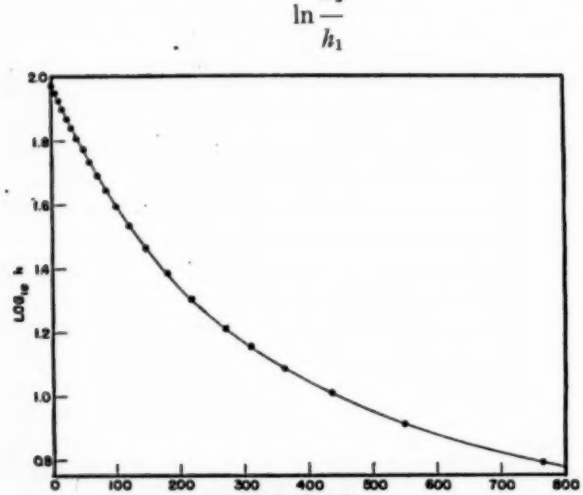


FIG. 3. Plot of $\log_{10}h$ *vs* time for a synthetic rubber latex at 30°C.

[3] G. Barr, *A Monograph of Viscometry* (Oxford University Press, New York, 1931).

as the effective head for Newtonian liquids. This mean value of h is usable for non-Newtonian fluids if $(h_2-h_1)/h_1\ll 1$. In terms of these the apparent fluidity ϕ_a becomes

$$\phi_a=\frac{8L\bar{Q}}{\pi R^4\alpha\bar{h}} \tag{25}$$

and the maximum shearing stress F_m,

$$F_m=\frac{RP}{2L}=\frac{R\alpha\bar{h}}{2L}. \tag{26}$$

Now, it was shown by Krieger and Maron[4] that the relation between F_m and the rate of shear $g(F_m)$ is given by

$$\frac{g(F_m)}{F_m}=\phi_a\left[1+\frac{1}{4}\frac{d\ln\phi_a}{d\ln F_m}\right]. \tag{27}$$

Consequently, construction of a plot of $\log_{10}\phi_a$ *vs* $\log_{10}\bar{h}$ and its differentiation at various values of $\bar{h}$ yields $d\log_{10}\phi_a/d\log_{10}\bar{h}$, which is identical with $d\ln\phi_a/d\ln F_m$. With the values of this quantity determined, $g(F_m)$ can be calculated for various values of F_m to obtain the flow curve.

In the second, or differential method, of treating the data F_m is

$$F_m=\frac{R\alpha h}{2L} \tag{28}$$

and ϕ_a from Eq. (22) is

$$\phi_a=-\frac{8(2.303)Lr^2m}{R^4\alpha}=-\frac{m}{B}. \tag{29}$$

[4] I. M. Krieger and S. H. Maron, J. Appl. Phys. **25**, 72 (1954).

TABLE II. Determination of rate of shear-shearing stress relation for 62.2 percent solids synthetic latex by differential method at 30°C.

Capillary D: $R=3.818\times10^{-2}$ cm, $L=14.088$ cm
Mercury manometer: $r_m=0.1674$ cm, $Rc=1.268$ cm

h cm	t sec	$-m\times10^5$	$\frac{dm}{dt}\times10^5$	Δ[a]	ϕ_a	$\frac{g(F_m)}{F_m}$	F_m dynes/cm^2	$g(F_m)$ sec^{-1}
94.25	0	4.605	1.799	0.0921	1.168	1.276	1723	2199
89.25	5.4	4.455	1.759	0.0962	1.130	1.239	1632	2022
84.25	10.8	4.400	1.734	0.0972	1.116	1.224	1541	1886
79.25	17.0	4.295	1.695	0.0998	1.090	1.199	1449	1737
74.25	23.8	4.178	1.684	0.1048	1.060	1.171	1358	1590
69.25	31.2	4.065	1.653	0.1086	1.031	1.143	1266	1447
64.25	39.2	3.925	1.589	0.1120	0.996	1.108	1175	1302
59.25	48.8	3.762	1.522	0.1167	0.954	1.065	1083	1153
54.25	58.2	3.638	1.464	0.1201	0.923	1.034	992.0	1026
49.25	70.8	3.440	1.389	0.1275	0.873	0.984	900.5	886.1
44.25	85.2	3.248	1.322	0.1362	0.824	0.936	809.1	757.3
39.25	101.4	3.022	1.246	0.1481	0.767	0.881	717.7	632.3
34.25	121.6	2.788	1.134	0.1585	0.707	0.819	626.3	512.9
29.25	147.0	2.562	1.016	0.1680	0.650	0.759	534.8	405.9
24.25	181.0	2.268	0.881	0.1861	0.575	0.682	443.4	302.4
20.25	218.2	1.940	0.751	0.2168	0.492	0.599	370.3	221.8
16.25	272.2	1.600	0.582	0.2409	0.406	0.504	297.1	149.7
14.25	310.8	1.425	0.480	0.2566	0.361	0.454	260.6	118.3
12.25	363.4	1.170	0.365	0.2896	0.297	0.383	224.0	85.8
10.25	436.0	0.970	0.264	0.3035	0.246	0.321	187.4	60.2
8.25	550.0	0.722	0.189	0.3904	0.183	0.254	150.9	38.3
6.25	765.0	0.418	0.101	0.6245	0.106	0.172	114.3	19.7

[a] $\Delta=\frac{1}{4}\frac{d\log\phi_a}{d\log F_m}$.

Substitution of these into Eq. (27) yields

$$\frac{g(F_m)}{F_m}=-\frac{m}{B}\left[1+\frac{1}{4(2.303)}\frac{\frac{dm}{dt}}{m^2}\right]=-\frac{m}{B}\left[1+\frac{1}{9.212m^2}\frac{dm}{dt}\right]. \quad (30)$$

In this method $\log_{10}h$ is plotted *vs* t and differentiated to give m. The m values are then plotted *vs* t and the plot again differentiated to give dm/dt. These derivatives are substituted into Eqs. (29) and (30) to find ϕ_a and $g(F_m)/F_m$. The steps in this procedure are illustrated by the data given in Table II, which were deduced from the measurements shown in Fig. 3.

Finally, Fig. 4 shows the flow curve of the latex as deduced from measurements in two capillaries of different radii. The fact that the two sets of data lie on a single curve confirms the theory of the instrument as applied to non-Newtonian fluids, and indicates that excellent results can be obtained with it.

Of the two methods discussed for obtaining the flow curves of non-Newtonian fluids, the second or differential method is exact, while the incremental one is approximate. When the deviation of a fluid from Newtonian behavior is not too pronounced, the two methods give identical results, and the incremental one is to be preferred since it involves only one graphical differentiation. However, when the fluid deviates appreciably from Newtonian flow, the differential method may be required to arrive at the relation between $g(F_m)$ and F_m.

CONCLUSIONS

The results of this paper show that the viscometer described is capable of giving excellent results when used with either Newtonian or non-Newtonian fluids. When used with Newtonian fluids, absolute viscosities may be measured with high accuracy without requiring a density determination. With non-Newtonian fluids the flow curve may be determined over a wide shearing-stress range from the observations made during a single run. Furthermore, by providing means for using lighter

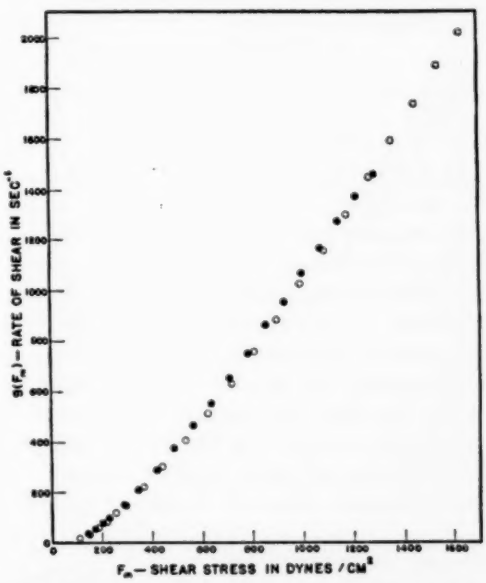


FIG. 4. Flow curve for 62.2 percent solids latex in two capillaries at 30°C.

liquids as well as mercury as drive fluids, the viscometer offers the means of investigating the low as well as the higher shearing-stress ranges, and to ascertain thereby more precisely whether yield points exist in the systems investigated.

ACKNOWLEDGMENT

This work was performed as a part of the research project sponsored by the Reconstruction Finance Corporation, Office of Synthetic Rubber, in connection with Government Synthetic Rubber Program.

APPENDIX

The only measurement made here which required a kinetic energy correction was that of water run with the mercury manometer. This correction was made as follows. Under conditions where the kinetic energy is significant, the pressure drop across the capillary is given by

$$P=\frac{8LQ\eta}{\pi R^4}+\frac{n\rho_s Q^2}{\pi^2 R^4} \tag{31}$$

where ρ_s is the density of the liquid measured and n is a constant which may be taken as 1.15. Substituting $P=\alpha h$ and $Q=-\pi r^2 dh/dt$, we get

$$\alpha h=-\frac{8Lr^2\eta}{R^4}\frac{dh}{dt}\left[1-\frac{n\rho_s r^2}{8L\eta}\frac{dh}{dt}\right]$$

which on integration yields

$$\left(\log_{10}h+\frac{n\alpha\rho_s R^4 h}{64(2.303)L^2\eta^2}\right)=-Bt+C \tag{32}$$

where B is given by Eq. (22) and C is a constant of integration. By estimating η from the low values of h, and using it along with the density, the quantity on the left of Eq. (32) can be evaluated and plotted *vs* t. Least squares calculation of the slope of the resulting straight line allows then the evaluation of η.

JOURNAL OF APPLIED PHYSICS VOLUME 25, NUMBER 8 AUGUST, 1954

Physical Theory of New Circuit Representation for Junction Transistors

JAKOB ZAWELS*
Radio Corporation of America, Victor Division, Camden, New Jersey
(Received November 9, 1953)

A circuit representation for a junction transistor is derived which places the mechanism of transistor action in evidence. The circuit is a direct interpretation of the diffusion equation and the boundary conditions which include the effects of base width modulation. It is found that the active part of the circuit is independent of frequency. A modification of the circuit for common emitter operation is shown and experimental results are given.

I. INTRODUCTION

THE physical properties of junction transistors have largely been described by Shockley,[1,2] Sparks,[2] and Teal.[2] The diffusion equation is the basic equation from which Shockley *et al.* derived some of the expressions for the electrical parameters. Early[3] extended this theory by stressing the importance of space charge layer widening at the collector junction and thereby accounted for the "collector resistance" of the T-equivalent circuit. However, it is possible to define the boundary conditions for the diffusion equation in a manner which will simultaneously account for all transistor parameters which have their basis in the diffusion process.

In either case, whether the Early effect is included or not, the use of the diffusion equation and the physical shape of a junction transistor implies a unique circuit in which each element may be associated with each individual phenomenon present. Obviously, this circuit could be transformed by various circuit manipulation into any number of circuits of arbitrary topology, such as a T, π, lattice, etc. However, such a transformed circuit suffers from two main disadvantages. Firstly, the elements lose their individual meaning (for example it is difficult to associate the role of the current generator in the T-equivalent circuit with an individual mechanism in the transistor). Secondly, a transformed circuit inevitably must employ gross approximations in order to limit its complexity.

In this work an equivalent circuit is derived which is a direct interpretation of the diffusion equation and the boundary conditions.

II. THE DIFFUSION EQUATION AND THE BOUNDARY CONDITIONS

Figure 1 shows a cross section of a plane parallel p-n-p junction transistor having a uniform base thick-

* Present address: Michaelandré House—Apt. 103, 133 Celliers Street, Pretoria, South Africa.
[1] W. Shockley, Bell System Tech. J. **28**, 435 (1949).
[2] Shockley, Sparks, and Teal, Phys. Rev. **83**, 151 (1951).
[3] J. M. Early, Proc. Inst. Radio Engrs. **40**, 1401 (1952).

ness W and connected to batteries V_E and V_C. The emitter and collector space charge regions are also shown. An X, Y coordinate system is indicated with its origin at the emitter junction. The behavior for holes injected into the base is described by [reference 1, Eq. (4.7)],

$$D_p\frac{\partial^2 p}{\partial x^2}+\frac{p_n-p}{\tau_p}=\frac{\partial p}{\partial t} \tag{1a}$$

where D_p is the diffusion constant, p is hole density, τ_p their lifetime, and p_n is the equilibrium value of holes in the n region.

(1) can be split into a dc and an ac equation by assuming a solution:

$$p(x,t)=P_0(x)+p_1(x)\exp(j\omega t)$$

which yields a dc equation

$$\frac{\partial^2 P_0}{\partial x^2}=\frac{P_0-p_n}{L_p^2} \tag{1b}$$

and an ac equation,

$$\frac{\partial^2 p_1}{\partial x^2}=\frac{p_1}{L_p^2}(1+j\omega\tau_p). \tag{1c}$$

The resemblance of the above equation to a transmission line differential equation is evident. However, the exact transmission line analogy will be established later.

The dc boundary conditions for P_0, that is the hole flow across the junction at $x=0$ and $x=W$, depend on the potential V_E and V_C across the junction.

Thus[1] at $x=0$,

$$P_0(0)=p_n e^{\Lambda V_E} \tag{2a}$$

and at $x=W$

$$P_0(W)=p_n e^{-\Lambda V_C} \tag{2b}$$

where

$$\Lambda\equiv q/kT.$$

Using these boundary conditions, the solution to (1b) is

$$P_0(x)=\frac{(e^{-\Lambda V_C}-e^{\Lambda V_E})}{2\sinh\left(\frac{W}{L_p}\right)}p_n\sinh\left(\frac{x-\frac{W}{2}}{L_p}\right)+\frac{(e^{-\Lambda V_C}+e^{\Lambda V_E}-2)}{2\cosh\left(\frac{W}{L_p}\right)}p_n\cosh\left(\frac{x-\frac{W}{2}}{L_p}\right)+p_n. \tag{3}$$

The boundary conditions for the solution of (1c) are given by a perturbation on the dc solution (3) as a result of a change in V_E or V_C.

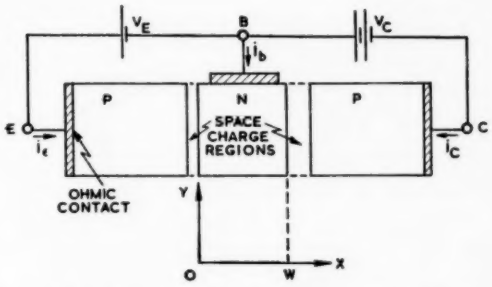


FIG. 1. Cross section of parallel plane junction transistor.

Thus, the ac boundary conditions are at $x=0$

$$p_1(0)=\frac{\partial P_0}{\partial V_E}v_e=p_n\Lambda e^{\Lambda V_E}v_e \tag{4a}$$

and at $x=W$,

$$p_1(W)=\frac{\partial P_0}{\partial V_C}v_c \tag{4b}$$

where v_e=small signal emitter voltage, v_c=small signal collector voltage.

In evaluating $p_1(W)$ we note that W is a function of V_C due to the Early effect. Furthermore, at $x=W$, P_0 [in Eq. (3)] virtually vanishes especially for collector voltage $V_C>\frac{1}{2}$ volt. Hence, the boundary condition (4b) is rewritten

$$p_1(W)=\left.\frac{\partial P_0(x)}{\partial x}\right|_{x=W}\cdot\frac{\partial W}{\partial V_C}v_c+\frac{\partial P_0(W)}{\partial V_C}v_c. \tag{4c}$$

From (3) and (4c)

$$p_1(W)=\frac{p_n\Lambda e^{\Lambda V_E}}{K}v_c \tag{4d}$$

where

$$\frac{1}{K}\equiv\left[\frac{1}{\Lambda L_p}\cdot\frac{\partial W}{\partial V_C}\operatorname{cosech}\frac{W}{L_p}-\left(\frac{e^{-\Lambda V_C}}{e^{\Lambda V_E}}\right)\frac{1}{\Lambda L_p}\frac{\partial W}{\partial V_C}\coth\frac{W}{L_p}+\frac{e^{-\Lambda V_C}}{e^{\Lambda V_E}}\right]. \tag{5}$$

The first term in (5) is much larger than the other two terms which involve $e^{-\Lambda V_C}$ for $V_C>0.5$ volt. Stating Eq. (5) in words: K is the ratio of the boundary values of the incremental hole densities, ($p_1(0)$ and $p_1(W)$), at the emitter and collector junctions, respectively.

Using boundary conditions (4a) and (4d) the solution to the ac equation (1c) may be found. Furthermore, it follows from the definition of the diffusion constant that the current due to the concentration gradient of holes is

$$i_p=-qD\frac{\partial p}{\partial x}.$$

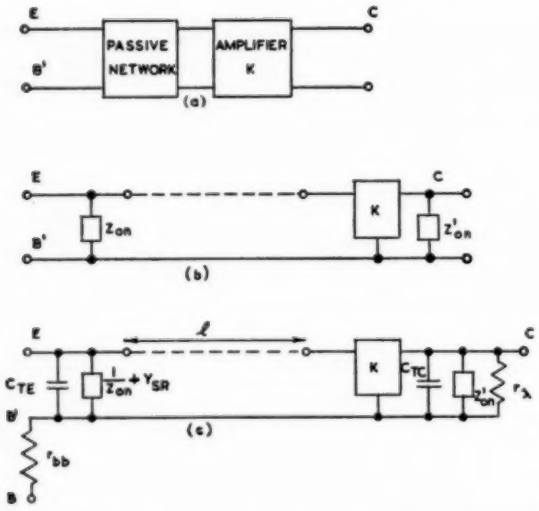


FIG. 2. (a) Block diagram of intrinsic transistor. (b) Equivalent circuit of intrinsic transistor. (c) Equivalent circuit of junction transistor.

Hence the currents due to holes at the emitter ($x=0$) may be written

$$i_e'=\frac{qp_n\mu_p}{L_p}e^{\Delta V_E}(1+j\omega\tau_p)^{\frac{1}{2}}\left\{v_e\coth\frac{W}{L_p}(1+j\omega\tau_p)^{\frac{1}{2}}-\frac{v_c}{K}\operatorname{cosech}\frac{W}{L_p}(1+j\omega\tau_p)^{\frac{1}{2}}\right\} \quad (6a)$$

and at the collector ($x=W$)

$$i_c'=\frac{qp_n\mu_p}{L_p}e^{\Delta V_E}(1+j\omega\tau_p)^{\frac{1}{2}}\left\{v_e\operatorname{cosech}\frac{W}{L_p}(1+j\omega\tau_p)^{\frac{1}{2}}-\frac{v_c}{K}\coth\frac{W}{L_p}(1+j\omega\tau_p)^{\frac{1}{2}}\right\}. \quad (6b)$$

The ac electron currents which flow in a p-n-p transistor are virtually unaffected by the base width modulation of the collector voltage. It may be written by inspection with the help of (6a) and (6b) as follows:

$$i_e''=\frac{qn_p\mu_n}{L_n}e^{\Delta V_E}(1+j\omega\tau_n)^{\frac{1}{2}}v_e \quad (7a)$$

and

$$i_c''=\frac{qn_p'\mu_n}{L_n'}e^{-\Delta V_C}(1+j\omega\tau_n')^{\frac{1}{2}}v_c. \quad (7b)$$

The primed quantities n_p' and τ_n' refer to the collector.

The total emitter and collector currents flowing now are the sums of (6a) and (7a) and also (7a) and (7b), respectively.

III. THE INTRINSIC TRANSISTOR

The node equations for a four terminal network are written generally.

$$i_1=y_{11}v_1+y_{12}v_2 \quad (8a)$$

$$i_2=y_{21}v_1+y_{22}v_2. \quad (8b)$$

(Currents flowing into positive terminals are considered positive.)

Identifying subscripts 1 and 2 with ϵ and c, and comparing coefficients with the total currents flowing, it is seen that

$$y_{11}=G\theta\coth\theta+G\phi \quad (9a)$$

$$y_{21}=-G\theta\operatorname{cosech}\theta \quad (9b)$$

$$y_{12}=-\frac{G\theta}{K}\operatorname{cosech}\theta \quad (9c)$$

$$y_{22}=\frac{G\theta}{K}\coth\theta+G\phi'\frac{e^{-\Delta V_C}}{e^{\Delta V_E}}, \quad (9d)$$

where

$$G\equiv\frac{q}{W}p_n\mu_pe^{\Delta V_E}\quad\left(\doteq\frac{qI_E}{kT}\right) \quad (10a)$$

$$\theta\equiv\frac{W}{L_p}(1+j\omega\tau_p)^{\frac{1}{2}} \quad (10b)$$

$$\phi\equiv\frac{D_nn_pW}{D_pp_nL_n}(1+j\omega\tau_n)^{\frac{1}{2}} \quad (10c)$$

$$\phi'\equiv\frac{D_nn_p'W}{D_pp_nL_n'}(1+j\omega\tau_n')^{\frac{1}{2}}. \quad (10d)$$

K, as defined in Eq. (5), is a frequency independent number and may be in the order of 10 000.

Consider Eqs. (9a) to (9d). For the moment, neglect the terms containing ϕ and ϕ'. Now it is seen that $y_{22}/K=y_{11}$ and $y_{12}/K=y_{21}$. This is interpreted to mean that there is present a symmetrical passive network (where $y_{11}=y_{22}$ and $y_{12}=y_{21}$) to which an impedance multiplier K is connected as shown in Fig. 2a. On closer examination it is found that this impedance multiplier which we will call a "K amplifier" has the properties:

$$\frac{\text{output voltage}}{\text{input voltage}}=K$$

and

$$\frac{\text{output current}}{\text{input current}}=\text{Unity}.$$

In fact the properties of a K amplifier are identical to those of a grounded grid (vacuum tube) triode whose plate resistance is zero and whose amplification factor $\mu=K-1\doteq K$.

Also, the passive network is an RC transmission line

having a characteristic impedance,

$$Z_0=\left(\frac{z}{y}\right)^{\frac{1}{2}}=\frac{1}{G\theta}=\frac{1}{\dfrac{W}{G\dfrac{W}{L_p}(1+j\omega\tau_p)^{\frac{1}{2}}}} \tag{11a}$$

and a hyperbolic angle

$$(zy)^{\frac{1}{2}}l=\theta=\frac{W}{L_p}(1+j\omega\tau_p)^{\frac{1}{2}}, \tag{11b}$$

where

$$z=r+j\omega L=\text{series impedance per unit length}, \tag{11c}$$

$$y=g+j\omega C=\text{shunt admittance per unit length}, \tag{11d}$$

and

$$l=\text{length of line}. \tag{11e}$$

Substituting (11c) and (11d) into (11a) and (11b), it follows that the transmission line has zero series inductance while its total series resistance $rl=1/G$, the total shunt conductance $gl=GW^2/L_p^2$ and the total shunt capacitance $Cl=GW^2\tau_p/L_p^2$.

So far, terms involving ϕ and ϕ' in Eqs. (9a) and (9d) have been neglected. This is easily accounted for by two impedances placed in parallel at each end of the circuit as seen in Fig. 2b. Here

$$\frac{1}{Z_{0n}}=G\phi=G\frac{D_n n_p W}{D_p p_n L_n}(1+j\omega\tau_n)^{\frac{1}{2}}.$$

This impedance Z_{0n} is, for a good transistor, very large compared to the characteristic impedance of the line Z_0. Also,

$$\frac{1}{Z'_{0n}}=G\phi'\frac{e^{-\Delta V_C}}{e^{\Delta V_E}}$$

and is usually entirely negligible compared to the other conductances which it shunts, especially for collector voltages greater than 0.5 volt. Figure 2b shall be referred to as the intrinsic transistor.

It is seen that the transmission line represents the passage of the minority carriers through the base. The K amplifier represents the falling of the minority carriers down the potential hill, thereby resulting in voltage, but no current, multiplication. Z_{0n} (and also Z_{0n}') represents the impedance presented to the flow of majority carriers from the base to the emitter (and also to the collector).

IV. PRACTICAL TRANSISTOR

In order to obtain the circuit representation for a practical transistor additional elements must be added to the intrinsic transistor as shown in Fig. 2c. The most important of these are the collector barrier, or transition capacitance[1] C_{TC}, and the base spreading resistance r_{bb} (which is due to the bulk resistivity of the germanium). A less important effect for a good transistor is r_λ, which is the collector leakage resistance due to the extraneous matter shunting the collector junction. Also, for present-day transistors, C_{TE}, the emitter barrier capacitance, is very small compared to the capacitance of the transmission line. Furthermore, should surface recombination exist (on the surface of the base region), as is often the case with alloy transistors,[4] an admittance Y_{SR} should be added probably in parallel with Z_{0n}.

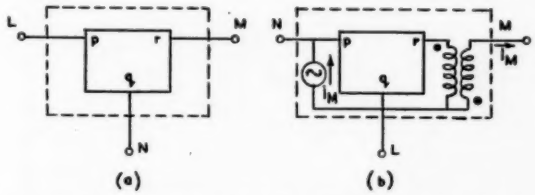


FIG. 3. Three-terminal device LMN represented by two circuits which are approximately equivalent if $y_{11}\gg y_{12}$ and $y_{21}\gg y_{22}$.

Finally, possible effects like the modulation of the base spreading resistance r_{bb}, by virtue of collector barrier width modulation,[3] and ohmic series resistance in the emitter and collector due to poor lead contact, or the material spreading resistance, have been omitted either because they are normally negligible or because they are experimentally difficult to distinguish from the other elements that are present.

V. GROUNDED EMITTER CIRCUIT

For the case where the emitter is the common lead (in some amplifier applications) difficulty may be encountered if Fig. 2c is used. Thus, we shall modify this circuit slightly by the use of the following circuit theorem without losing the physical meaning in the resultant circuit.

Consider a network pqr as seen in Figs. 3a and 3b. Let a phase reversing transformer and current generator i_M ne connected to pqr as seen in Fig. 3b. Now, the electrical characteristics as viewed from terminals L, M, and N are virtually identical for the two devices shown in Figs. 3a and 3b, providing that the network pqr has the following properties

$$y_{11}\gg y_{12} \quad \text{and} \quad y_{21}\gg y_{22} \tag{12}$$

where p and q are regarded as the input leads, while r and q are the output leads.

It now follows that a device may be represented by two circuits which differ only in the respect that one of them employs a phase reversing transformer and a current generator, whose current is proportional to the current in the output lead, providing that the inequality (12) holds. Applying this theorem: Fig. 2a represents a grounded base circuit, while Fig. 4a represents a grounded emitter circuit; both employ identically the same passive network and K amplifier.

It is obvious that the sequence of the placing of the passive network, reversing transformer, and K ampli-

[4] A. R. Moore and J. I. Pankove (private communication).

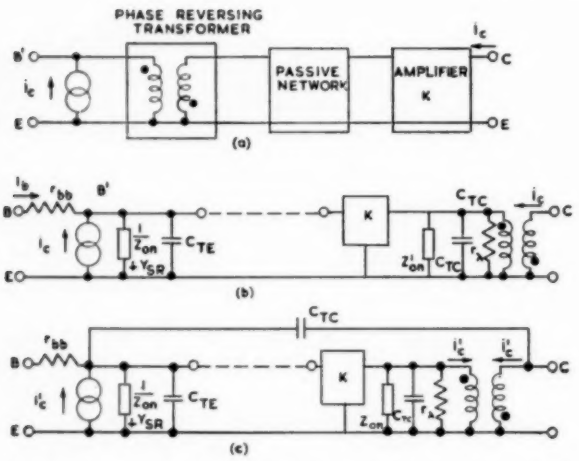


FIG. 4. (a) Common emitter block diagram. (b) Equivalent circuit for common emitter connection. (c) Alternate representation of circuit shown in Fig. 4b.

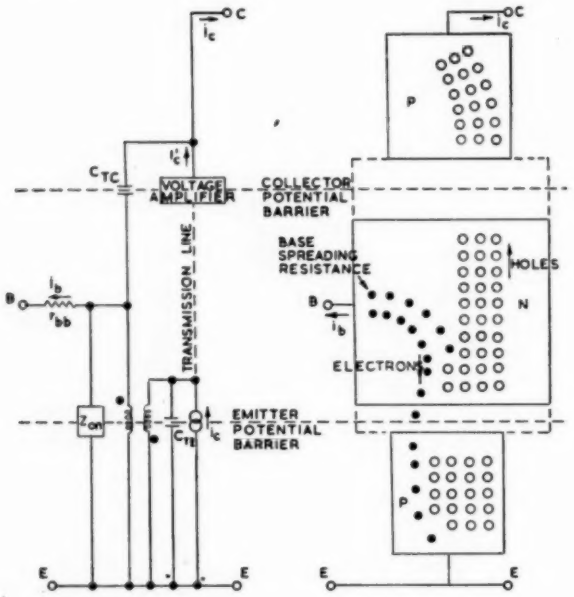


FIG. 4d. Pictorial association.

fier is immaterial provided that the impedance transforming properties of the amplifier are recognized. This is a result of these three sections being simply in cascade.

Since the inequality (12) holds for both the intrinsic transistor (Fig. 2b) as well as the practical transistor (Fig. 2c), there is a freedom of choice as to where the external elements of the practical transistor should now be connected. Thus, C_{TC} which is connected in parallel with the amplifier, between point C and B' in Fig. 2c, may be connected either in parallel with the K amplifier as shown in Fig. 4b or between point C and B' as shown in Fig. 4c, depending on whether the circuit theorem is applied to Fig. 2c or Fig. 2b, respectively. Similarly, there are two choices for the placing of r_{bb}, one of which is shown.

The sequence of the elements may also be rearranged as shown on the left side of Fig. 4d. Now, the flow of holes and electrons may be traced in both the pictorial representation of the transistor and the equivalent circuit, and to a large extent a one-to-one correspondence can be established between each element on the left and a particular phenomenon to be found in the actual transistor on the right.

VI. EXPERIMENTAL

In order to determine the elements of a four terminal device, four independent measurements as a function of frequency are required. Those shown in Fig. 5 are for the common emitter connection of an alloy junction transistor.

The magnitude of the input impedance and its phase angle for a short circuited output is shown in Fig. 5a. The magnitude of the short-circuit current amplification factor and its phase angle is shown in Fig. 5b. The equivalent parallel output resistance and capacitance for an ac open-circuit input is shown in Fig. 5c. The magnitude of the reverse voltage amplification factor, and its phase angle, for an ac open circuit at the input is shown in Fig. 5d. The small circles are experimental points while the solid curves are calculated[5] after the hyperbolic functions of Eqs. 9(a) to 9(d) are approximated,[6] relying on the fact that for a useful transistor $W^2/L_p^2 \ll 1$.

For this transistor it was found that at the higher frequencies a small capacitance C_f of 4.4 micromicrofarad existed between the collector and the base, probably due to socket capacitance as well as due to the distributed nature of the base material in the vicinity of the collector transition capacitance. In Fig. 5d the solid line includes the effect of C_f in the calculated curve while the dotted line excludes it.

From Figs. 5a to 5d it is found that the first-order elements of Fig. 2a (or Fig. 4b) are: $r_{bb}=310$ ohm, $C_{TC}=49$ mmf; $r_\lambda=24$ megohm; $K=8500$; $rl=15$ ohm; $1/gl=255$ ohm, $Cl=0.042$ microfarad, where r, g, and C are the resistance, conductance and capacitance per unit length and l is the length of the transmission line.

VII. CONCLUSION

A circuit representation of a junction transistor has been derived which places in evidence the mechanism of transistor action. Thus for example, it is seen that the passage of minority carriers through the base and their drop through the collector potential barrier without multiplication can intuitively be associated with

[5] The curves are calculated by assuming that the calculated and experimental value at 1 kc coincided. Hence, the curves aim primarily to show the correctness of the elements used in the equivalent circuit and their relative values. This assumption (at 1 kc) is justified by the closeness of the magnitude of the physical constants, such as base width which this implies, to their estimated values.

[6] The degree of approximation is well within the experimental error.

the flow of current through the transmission line and the K amplifier, respectively.

Furthermore, it is found that the active part of the circuit, the K amplifier, is independent of frequency, and a variation of the circuit, which has the emitter as a common element employs a current generator which is also independent of frequency.

It is also noted that even in the absence of any modulation of the base width by the collector voltage, the circuit is unaltered except for the number K.

It is easily seen that the separation of the active and passive part and their cascaded layout simplifies the employment of the circuit in the calculation of gain, impedance levels, etc. Obviously, the transmission line may be approximated by lumped elements. In fact, the complexity of the circuit can easily be altered to meet the degree of accuracy required for a particular calculation. In its complete form the circuit represents the diffusion equation and the boundary conditions without approximation.

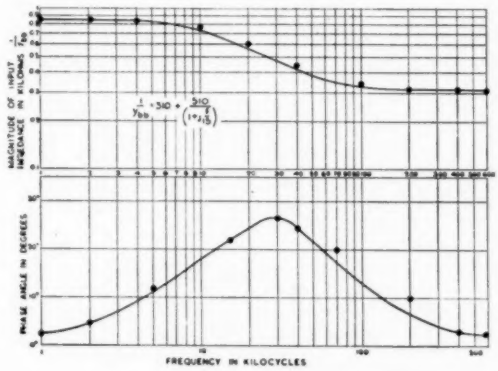


FIG. 5a. Input impedance $1/y_{bb}$.

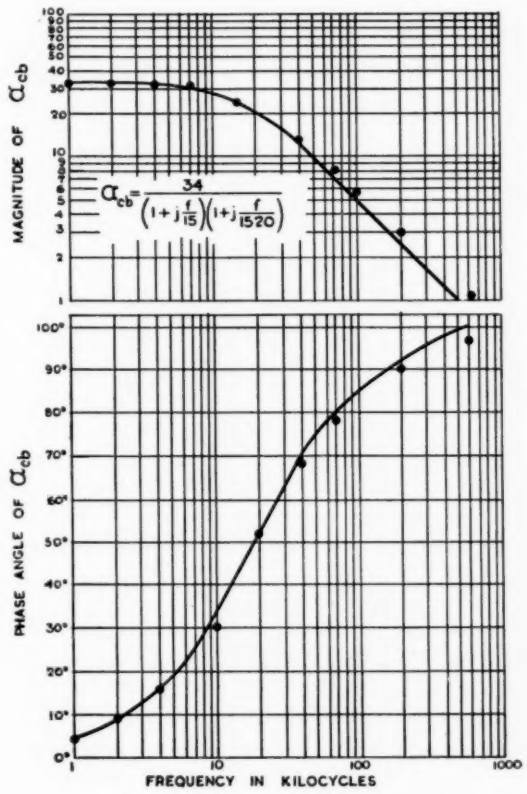


FIG. 5b. Short-circuit current amplification α_{cb}.

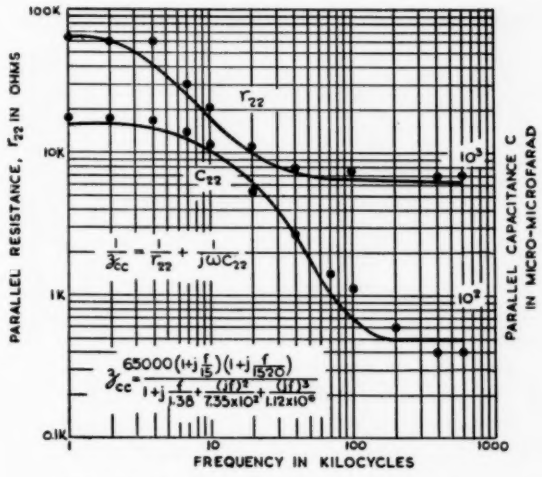


FIG. 5c. Open-circuit output impedance z_{ce}.

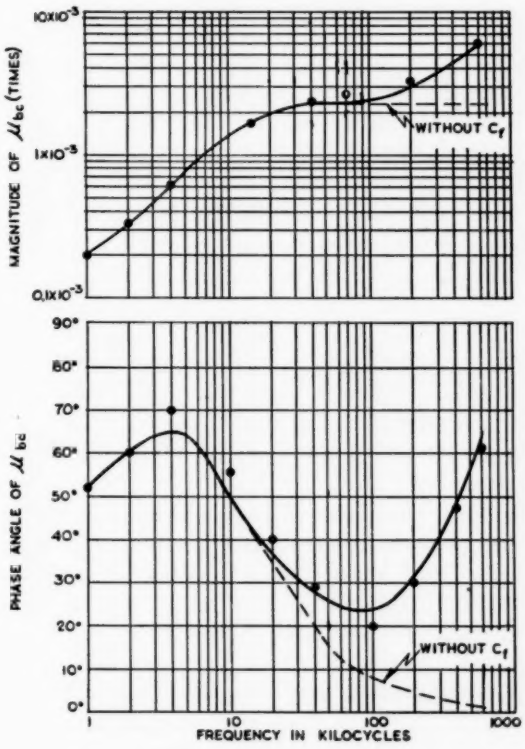


FIG. 5d. Reverse voltage amplification μ_{bc}.

Experimental Procedures for Determining the Efficiency of Four-Terminal Networks*

H. F. MATHIS
School of Electrical Engineering, University of Oklahoma, Norman, Oklahoma
(Received November 6, 1953)

Five experimental procedures which may be used for determining the efficiency of four-terminal networks, by means of circular loci plotted in the impedance or reflection coefficient planes, are presented.

INTRODUCTION

LIMITED discussions of this subject have been written by Barlow and Cullen,[1] Beatty,[2] King,[3] and Macalpine.[4] However, none of these references present methods for determining the efficiency of four-terminal networks which are satisfactory for all types of networks.

The procedures which are presented here were developed from well-known four-terminal circuit theory. No simplifying assumptions were made. Either an impedance bridge or a slotted line can be used for obtaining data. Graphical methods are used for handling the data. The theoretical basis for these procedures is given in the Appendix.

The arbitrary four-terminal network shown in Fig. 1 will be considered. Phasors will be denoted by bold face type, e.g. $\mathbf{E}$. The magnitudes will be denoted by either E or $|\mathbf{E}|$.

PROCEDURE I

This procedure can be used when an impedance bridge is available and it is convenient to replace the load impedance $\mathbf{Z}_L$ by at least three different values of reactance. It is not necessary to know the values of the reactances.

The load impedance $\mathbf{Z}_L$ is connected to the four-terminal network and the input impedance $\mathbf{Z}_A$ is measured. The load impedance is replaced by a reactance and the input impedance $\mathbf{Z}_x$ is measured. This is continued until at least three different values of $\mathbf{Z}_x$ are obtained.

The normalized values of $\mathbf{Z}_A$ and $\mathbf{Z}_x$ are plotted on a Smith chart,[5] or a $Z-\theta$ chart, as shown in Fig. 2. These impedances can be normalized with respect to any convenient value of R_0. A circle K is drawn through the values of $\mathbf{Z}_x$. A straight line is drawn through $\mathbf{Z}_A$ and the center of the circle K. The efficiency η is given by the formula

$$\eta=\frac{2D_1D_2}{(D_1+D_2)(1-D_3^2)}, \tag{1}$$

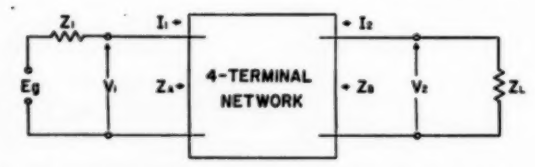


FIG. 1. A four-terminal network with external connections.

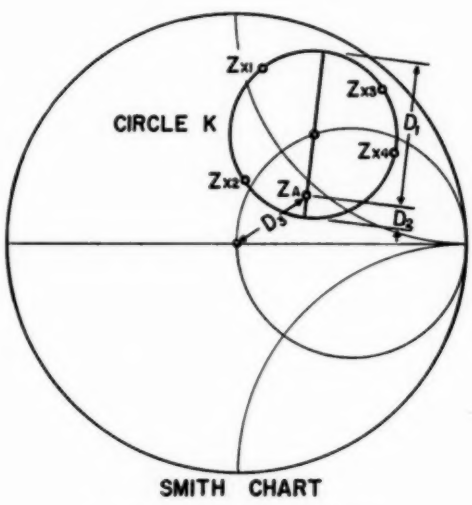


FIG. 2. Diagram for Procedure I.

where D_1, D_2, and D_3 are measured as fractions of the radius of the chart.

PROCEDURE II

This procedure can be used when a slotted line is available and it is convenient to replace the load impedance $\mathbf{Z}_L$ by at least three different values of reactance. It is not necessary to know the values of the reactances.

The load impedance $\mathbf{Z}_L$ is connected to the four-terminal network, the VSWR in the slotted line feeding the network is measured, and the position of a voltage minimum is located. The load impedance is replaced by a reactance, the VSWR is measured, the position of a voltage minimum is located, and the distance l from the voltage minimum for $\mathbf{Z}_L$ to the voltage minimum for the reactance is measured. This procedure is continued until at least three different reactances are con-

* This paper was written while the author was employed on a research project sponsored by the U. S. Air Force under contract No. AF33 (038)-10405, Supplementary Agreement 56 (53-606).

[1] H. M. Barlow and A. L. Cullen, *Micro-Wave Measurements* (Constable and Company, London, 1950), pp. 216-223.

[2] R. W. Beatty, Proc. Inst. Radio Engrs., **38**, No. 8, 893-7 (1950).

[3] D. D. King, *Measurements at Centimeter Wavelength* (D. Van Nostrand Company, Inc., New York, 1952), pp. 99-105.

[4] W. W. Macalpine, Elec. Eng. **72**, No. 10, 868 (1953).

[5] P. H. Smith, Electronics **12**, 29 (1939); **17**, 130 (1944).

nected to the network. It is necessary to assume a positive direction for l. The wavelength λ_g in the slotted line is either measured or calculated.

The values of $\mathbf{K}_L$ and $\mathbf{K}_x$ given by the formulas

$$\mathbf{K}_L=\frac{\text{VSWR}-1}{\text{VSWR}+1}\,\underline{|0^0} \tag{2}$$

and

$$\mathbf{K}_x=\frac{\text{VSWR}-1}{\text{VSWR}+1}\,\underline{|(l/\lambda_g)720^0} \tag{3}$$

are plotted on polar graph paper as shown in Fig. 3. A circle K is drawn through the values of $\mathbf{K}_x$. A straight line is drawn through $\mathbf{K}_L$ and the center of the circle K. The efficiency is given by Eq. (1).

PROCEDURE III

This procedure can be used when an impedance bridge is available and it is convenient to replace the load impedance $\mathbf{Z}_L$ by a short circuit and an open circuit. First the load impedance $\mathbf{Z}_L$ is connected to the bridge and its power factor $\cos\theta$ is determined.

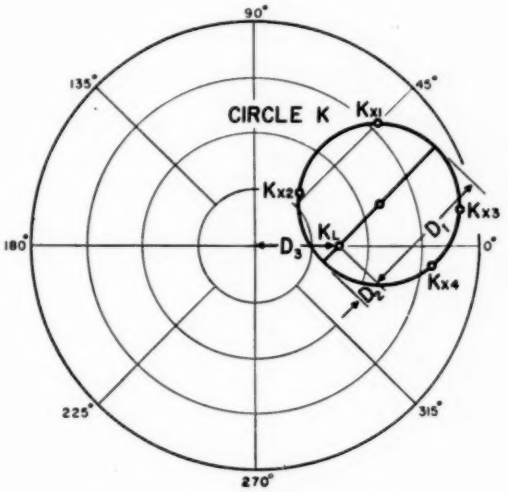


FIG. 3. Diagram for Procedure II.

Next $\mathbf{Z}_L$ is connected to the output terminals of the four-terminal network and the input impedance $\mathbf{Z}_A$ is measured. Then $\mathbf{Z}_L$ is replaced by a short circuit and the input impedance $\mathbf{Z}_{sc}$ to the network is measured. Finally, the short circuit is replaced by an open circuit and the input impedance $\mathbf{Z}_{oc}$ is measured.

The normalized values of $\mathbf{Z}_A$, $\mathbf{Z}_{sc}$, and $\mathbf{Z}_{oc}$ are plotted on a Smith chart, or a $Z-\theta$ chart, as show in Fig. 4. The efficiency is given by the formula

$$\eta=\frac{2d_1d_2\cos\theta}{d_3(1-d_4^2)}, \tag{4}$$

where d_1, d_2, d_3, and d_4 are measured as fractions of the radius of the chart.

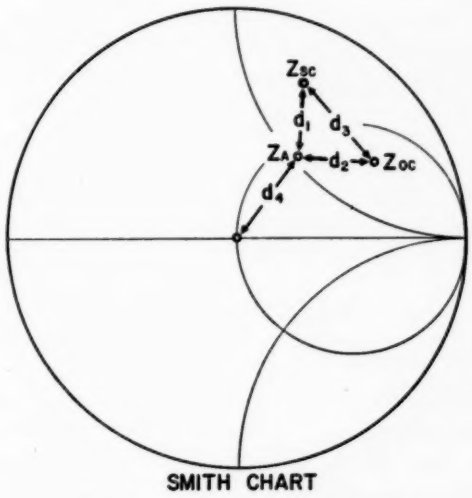


FIG. 4. Diagram for Procedure III.

PROCEDURE IV

This procedure can be used when a slotted line is available and it is convenient to replace the load impedance $\mathbf{Z}_L$ by a short circuit and an open circuit. First the power factor $\cos\theta$ of $\mathbf{Z}_L$ is determined.

If $\mathbf{Z}_L$ is connected to the slotted line, the value of $\mathbf{Z}_L$ is given by the equation

$$\frac{\mathbf{Z}_L-R_0}{\mathbf{Z}_L+R_0}=\frac{\text{VSWR}-1}{\text{VSWR}+1}\,\underline{|180^0+(s/\lambda_g)720^\circ}, \tag{5}$$

where R_0 is the characteristic impedance of the slotted line, s is the distance from $\mathbf{Z}_L$ to a voltage minimum on the slotted, and λ_g is the wavelength in the slotted line. It can be shown that

$$\cos\theta=\frac{1-A}{\sqrt{(1-A)^2+4A\,\sin^2(s/\lambda_g)720^\circ}}, \tag{6}$$

where

$$A=\left(\frac{\text{VSWR}-1}{\text{VSWR}+1}\right)^2. \tag{7}$$

When $\mathbf{Z}_L$ is replaced by a short circuit, the voltage minima are located at points which are $n\lambda_g/2$ from the short circuit, where n is an integer. When $\mathbf{Z}_L$ is replaced by an open circuit, the voltage minima are $(n+\frac{1}{2})\lambda_g/2$ from the open circuit. Consequently, Eq. (6) is also valid if s is defined as the distance from a voltage minimum when $\mathbf{Z}_L$ is connected to the slotted line to a voltage minimum when either a short or an open circuit replaces $\mathbf{Z}_L$.

The load impedance $\mathbf{Z}_L$ is connected to the slotted line, the VSWR is measured, and the position of a voltage minimum is located. Then $\mathbf{Z}_L$ is replaced by either a short circuit or an open circuit and a voltage minimum is located. The distance s between the two voltage minima is measured and the value of $\cos\theta$ is calculated using Eq. (6).

The load impedance $\mathbf{Z}_L$ is connected to the four-terminal network, the VSWR is the slotted line feeding the network is measured, and the position of a voltage minimum is located. The load impedance is replaced by a short circuit, the VSWR is measured, the position of a voltage minimum is located, and the distance l from the voltage minimum for $\mathbf{Z}_L$ to the voltage minimum for the circuit is measured. This procedure is repeated for an open circuit. It is necessary to assume a positive direction for l.

The values of $\mathbf{K}_L$, $\mathbf{K}_{sc}$, and $\mathbf{K}_{oc}$, given by Eqs. (2) and (3), are plotted on polar graph paper as shown in Fig. 5. The efficiency is given by Eq. (4).

When $\cos\theta \ll 1$, the losses in the connections between the slotted line and $\mathbf{Z}_L$ may not be negligible. If δ denotes the efficiency of power transfer through the connections

$$\mathrm{VSWR}_{\mathrm{obs}}=(1+\delta K_0)/(1-\delta K_0), \tag{8}$$

where

$$\mathbf{K}_0=(\mathbf{Z}_L-R_0)/(\mathbf{Z}_L+R_0). \tag{9}$$

If $\mathbf{Z}_L$ is replaced by either a short circuit or an open circuit,

$$\mathrm{VSWR}_\delta=(1+\delta)/(1-\delta). \tag{10}$$

The corrected VSWR is

$$\mathrm{VSWR}_{\mathrm{corrected}}=\frac{\mathrm{VSWR}_\delta\mathrm{VSWR}_{\mathrm{obs}}-1}{\mathrm{VSWR}_\delta-\mathrm{VSWR}_{\mathrm{obs}}}. \tag{11}$$

The corrected value of VSWR should be used in Eq. (7).

Also when $\cos\theta \ll 1$, $\mathrm{VSWR}_{\mathrm{obs}}$ and VSWR_δ will be very large and special methods must be used to determine their values. Either of the methods described by Terman and Pettit[6] should be satisfactory.

PROCEDURE V

For some networks, it may be found that the above procedures give a value for the efficiency of approximately 0/0. In such a case, it may be desirable to use an auxiliary network. The auxiliary network is connected as shown in Fig. 6.

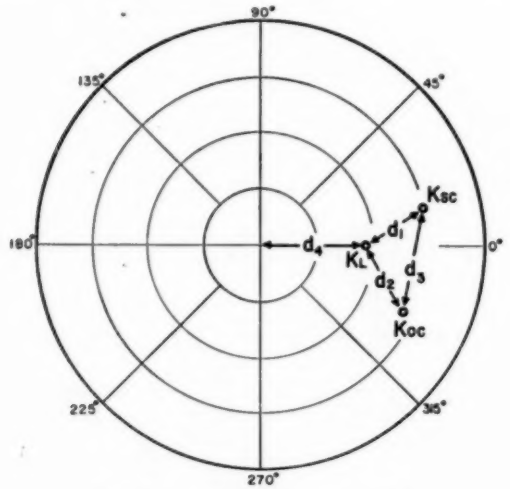


FIG. 5. Diagram for Procedure IV.

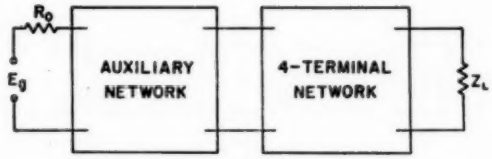


FIG. 6. Connections for Procedure V.

The overall efficiency η_{12} can be determined by any of the above procedures. Then the efficiency η_1 of the auxiliary network can be determined by any of the procedures. The efficiency η_2 of the four-terminal network is simply

$$\eta_2=\eta_{12}/\eta_1. \tag{12}$$

It may be observed that Procedures I–IV are much simpler when $\mathbf{Z}_A=R_0$, i.e., when $\mathbf{K}_L=0$. For this reason, it is desirable that the auxiliary network be a matching network. It is essential that the auxiliary network have a high efficiency.

APPENDIX: GENERAL THEORY

It is easily seen that, for the arbitrary four-terminal network shown in Fig. 1,

$$\mathbf{I}_1=\mathbf{E}_g/(\mathbf{Z}_1+\mathbf{Z}_A), \tag{13}$$

$$\mathbf{V}_1=\mathbf{I}_1\mathbf{Z}_A=\mathbf{E}_g\mathbf{Z}_A/(\mathbf{Z}_1+\mathbf{Z}_A), \tag{14}$$

and the power P_1 delivered to the network is

$$P_1=I_1^2R_A=E_g^2R_A/|\mathbf{Z}_1+\mathbf{Z}_A|^2, \tag{15}$$

where R_A is the resistive component of $\mathbf{Z}_A$.

According to well-known four-terminal circuit theory,

$$\mathbf{V}_1=\mathbf{A}\mathbf{V}_2-\mathbf{B}\mathbf{I}_2, \tag{16}$$

$$\mathbf{I}_1=\mathbf{C}\mathbf{V}_2-\mathbf{D}\mathbf{I}_2, \tag{17}$$

$$\mathbf{V}_2=\mathbf{D}\mathbf{V}_1-\mathbf{B}\mathbf{I}_1, \tag{18}$$

$$\mathbf{I}_2=\mathbf{C}\mathbf{V}_1-\mathbf{A}\mathbf{I}_1, \tag{19}$$

with

$$\mathbf{A}\mathbf{D}-\mathbf{B}\mathbf{C}=1. \tag{20}$$

It follows from Eqs. (13), (14), and (19) that

$$\mathbf{I}_2=\mathbf{E}_g(\mathbf{C}\mathbf{Z}_A-\mathbf{A})/(\mathbf{Z}_1+\mathbf{Z}_A). \tag{21}$$

The power P_2 delivered to the load impedance $\mathbf{Z}_L$ is

$$P_2=I_2^2R_L=E_g^2R_L|\mathbf{C}\mathbf{Z}_A-\mathbf{A}|^2/|\mathbf{Z}_1+\mathbf{Z}_A|^2, \tag{22}$$

where R_L is the resistive component of $\mathbf{Z}_L$. The efficiency η of power transfer is

$$\eta=P_2/P_1=|\mathbf{C}\mathbf{Z}_A-\mathbf{A}|^2R_L/R_A. \tag{23}$$

It should be noted that the efficiency is independent of $\mathbf{Z}_1$.

Now it will be assumed that the network is fed by a lossless transmission line and the signal generator is

[6] F. E. Terman and J. M. Pettit, *Electronic Measurements* (McGraw-Hill Book Company, Inc., New York, 1952), p. 140.

matched to the transmission line. The signal source and the lossless transmission line feeding the network can be represented by a simple generator having a generated voltage $\mathbf{E}_g$ and an internal impedance R_0. This assumption does not affect the efficiency of the network because it can be concluded from Eq. (23) that the efficiency is independent of how the network is fed.

It follows from Eqs. (16)–(19) that

$$\mathbf{Z}_A=(\mathbf{A}\mathbf{Z}_L+\mathbf{B})/(\mathbf{C}\mathbf{Z}_L+\mathbf{D}) \tag{24}$$

and

$$\mathbf{Z}_B=(\mathbf{D}R_0+\mathbf{B})/(\mathbf{C}R_0+\mathbf{A}). \tag{25}$$

When Eq. (24) is solved for $\mathbf{Z}_L$, the result is

$$\mathbf{Z}_L=-(\mathbf{D}\mathbf{Z}_A-\mathbf{B})/(\mathbf{C}\mathbf{Z}_A-\mathbf{A}). \tag{26}$$

If $\mathbf{Z}_L$ is replaced by an arbitrary reactance X, then the input impedance $\mathbf{Z}_x$ to the network is

$$\mathbf{Z}_x=(j\mathbf{A}X+\mathbf{B})/(j\mathbf{C}X+\mathbf{D}) \tag{27}$$

and the voltage reflection coefficient $\mathbf{K}_x$ is

$$\mathbf{K}_x=(\mathbf{Z}_x-R_0)/(\mathbf{Z}_x+R_0)$$
$$=-\frac{\mathbf{D}R_0-\mathbf{B}+j(\mathbf{C}R_0-\mathbf{A})X}{\mathbf{D}R_0+\mathbf{B}+j(\mathbf{C}R_0+\mathbf{A})X}. \tag{28}$$

Equation (28) can be written

$$\mathbf{K}_x=\frac{\mathbf{Z}_A-R_0}{\mathbf{Z}_A+R_0}-\frac{\mathbf{Z}_A-R_0}{\mathbf{Z}_A+R_0}\frac{\mathbf{D}R_0+\mathbf{B}+j(\mathbf{C}R_0+\mathbf{A})X}{\mathbf{D}R_0+\mathbf{B}+j(\mathbf{C}R_0+\mathbf{A})X}$$
$$-\frac{\mathbf{Z}_A+R_0}{\mathbf{Z}_A+R_0}\frac{\mathbf{D}R_0-\mathbf{B}+j(\mathbf{C}R_0-\mathbf{A})X}{\mathbf{D}R_0+\mathbf{B}+j(\mathbf{C}R_0+\mathbf{A})X}$$
$$=\frac{\mathbf{Z}_A-R_0}{\mathbf{Z}_A+R_0}-\frac{2R_0[\mathbf{D}\mathbf{Z}_A-\mathbf{B}+j(\mathbf{C}\mathbf{Z}_A-\mathbf{A})X]}{(\mathbf{Z}_A+R_0)[\mathbf{D}R_0+\mathbf{B}+j(\mathbf{C}R_0+\mathbf{A})X]}. \tag{29}$$

When the values of $(\mathbf{D}R_0+B)$ and $(\mathbf{D}\mathbf{Z}_A-\mathbf{B})$ given by Eqs. (25) and (26), respectively, are substituted in Eq. (29), the result is

$$\mathbf{K}_x=\frac{\mathbf{Z}_A-R_0}{\mathbf{Z}_A+R_0}+\frac{2R_0}{R_0+\mathbf{Z}_A}\frac{\mathbf{C}\mathbf{Z}_A-\mathbf{A}}{\mathbf{C}R_0+\mathbf{A}}\frac{\mathbf{Z}_L-jX}{\mathbf{Z}_B+jX}. \tag{30}$$

Since

$$\mathbf{Z}_L+\mathbf{Z}_B=-\frac{R_0+\mathbf{Z}_A}{(\mathbf{C}\mathbf{Z}_A-\mathbf{A})(\mathbf{C}R_0+\mathbf{A})}, \tag{31}$$

$$\mathbf{C}R_0+\mathbf{A}=-\frac{R_0+\mathbf{Z}_A}{(\mathbf{Z}_L+\mathbf{Z}_B)(\mathbf{C}\mathbf{Z}_A-\mathbf{A})} \tag{32}$$

and

$$\mathbf{K}_x=\frac{\mathbf{Z}_A-R_0}{\mathbf{Z}_A+R_0}-\frac{2R_0(\mathbf{Z}_L+\mathbf{Z}_B)(\mathbf{C}\mathbf{Z}_A-\mathbf{A})^2}{(R_0+\mathbf{Z}_A)^2}\frac{\mathbf{Z}_L-jX}{\mathbf{Z}_B+jX}. \tag{33}$$

For convenience in notation, let

$$\mathbf{K}_L=(\mathbf{Z}_A-R_0)/(\mathbf{Z}_A+R_0) \tag{34}$$

and

$$\mathbf{K}_x'=\frac{2R_0(\mathbf{Z}_L+\mathbf{Z}_B)(\mathbf{C}\mathbf{Z}_A-\mathbf{A})^2}{(R_0+\mathbf{Z}_A)^2}\frac{\mathbf{Z}_L-jX}{\mathbf{Z}_B+jX}. \tag{35}$$

If one lets

$$jX=(1+\mathbf{k})/(1-\mathbf{k}), \tag{36}$$

where $\mathbf{k}$ is an arbitrary complex number such that $k=1$, then

$$\mathbf{K}_x=\mathbf{K}_L-\frac{2R_0(\mathbf{Z}_L+\mathbf{Z}_B)(\mathbf{C}\mathbf{Z}_A-\mathbf{A})^2}{(R_0+\mathbf{Z}_A)^2}$$
$$\times\frac{\mathbf{Z}_L-1-(\mathbf{Z}_L+1)\mathbf{k}}{\mathbf{Z}_B+1-(\mathbf{Z}_B-1)\mathbf{k}}. \tag{37}$$

According to a well-known conformal mapping theorem, $\mathbf{K}_x$ will lie on a circle K when plotted on the complex plane. When the normalized values of $\mathbf{Z}_A$ and $\mathbf{Z}_x$ are plotted on a Smith chart, or a $Z-\theta$ chart, this is equivalent to plotting $\mathbf{K}_L$ and $\mathbf{K}_x$, respectively, on the complex plane. Thus as X varies, the locus of the normalized value of $\mathbf{Z}_x$ on a Smith chart, or a $Z-\theta$ chart, will be a circle K. The circle K must enclose the normalized value of $\mathbf{Z}_A$. If $\mathbf{Z}_B$ and $\mathbf{Z}_L$ are conjugate impedances, the center of the circle K and the point corresponding to $\mathbf{Z}_A$ will coincide.

Let X_1 and X_2 denote the values of X for which $\mathbf{K}_x'$ is a maximum and a minimum, respectively. Let $\mathbf{K}_1'$, $\mathbf{K}_2'$, $\mathbf{K}_1$, $\mathbf{K}_2$, $\mathbf{Z}_{x1}$, and $\mathbf{Z}_{x2}$ denote the values of $\mathbf{K}_x'$, $\mathbf{K}_x$, and $\mathbf{Z}_x$ corresponding to X_1 and X_2, respectively. After $\mathbf{K}_1$ and $\mathbf{K}_2'$ are determined by the usual analytical method for finding the maxima and minima of a function, it can be shown that

$$\frac{K_1'K_2'}{|\mathbf{K}_1'-\mathbf{K}_2'|}=\frac{2R_0R_L|\mathbf{C}\mathbf{Z}_A-\mathbf{A}|^2}{|R_0+\mathbf{Z}_A|^2}. \tag{38}$$

It can also be shown that

$$|R_0+\mathbf{Z}_A|^2=4R_0R_A/(1-K_L^2). \tag{39}$$

Consequently,

$$\frac{K_1'K_2'}{|\mathbf{K}_1'-\mathbf{K}_2'|}=R_L(1-K_L^2)|\mathbf{C}\mathbf{Z}_A-\mathbf{A}|^2/2R_A. \tag{40}$$

The derivation of Eq. (40) is not given in detail because of its length. If Eq. (40) is compared with Eq. (23), it is seen that

$$\eta=2K_1'K_2'/(1-K_L^2)|\mathbf{K}_1'-\mathbf{K}_2'|. \tag{41}$$

If the values of $\mathbf{K}_x'$ are plotted on the complex plane, their locus is a circle. It is easily seen that $\mathbf{K}_1'$ and $\mathbf{K}_2'$ can be found by drawing a straight line through the origin and the center of the circle. The points of intersection of this line and the circle are $\mathbf{K}_1'$ and $\mathbf{K}_2'$. Also, if $\mathbf{K}_x$ is plotted on the complex plane, the locus is the circle K. The values of $\mathbf{K}_1$ and $\mathbf{K}_2$ can be found by

drawing a straight line through $\mathbf{K}_L$ and the center of the circle. The points of intersection of the line and the circle are $\mathbf{K}_1$ and $\mathbf{K}_2$.

If $\mathbf{K}_{sc}'$ and $\mathbf{K}_{sc}$ denote the values of $\mathbf{K}_z'$ and $\mathbf{K}_z$, respectively, when $X=0$, then

$$\mathbf{K}_{sc}'=\mathbf{K}_{sc}-\mathbf{K}_L = -\frac{2R_0(\mathbf{Z}_L+\mathbf{Z}_B)(\mathbf{CZ}_A-\mathbf{A})^2}{(R_0+\mathbf{Z}_A)^2}\frac{\mathbf{Z}_L}{\mathbf{Z}_B}. \quad (42)$$

If $\mathbf{K}_{0c}'$ and $\mathbf{K}_{0c}$ denote the values of $\mathbf{K}_z'$ and $\mathbf{K}_z$, respectively, when $X=\infty$, then

$$\mathbf{K}_{0c}'=\mathbf{K}_{0c}-\mathbf{K}_L = \frac{2R_0(\mathbf{Z}_L+\mathbf{Z}_B)(\mathbf{CZ}_A-\mathbf{A})^2}{(R_0+\mathbf{Z}_A)^2}. \quad (43)$$

It is easily seen that

$$\frac{\mathbf{K}_{sc}'\mathbf{K}_{0c}'}{\mathbf{K}_{sc}'-\mathbf{K}_{0c}'}=\frac{2R_0\mathbf{Z}_L(\mathbf{CZ}_A-\mathbf{A})^2}{(R_0+\mathbf{Z}_A)^2}. \quad (44)$$

When Eq. (44) is compared with Eqs. (39) and (23), it is seen that

$$\eta=\frac{2K_{sc}'K_{0c}'\cos\theta}{|\mathbf{K}_{sc}'-\mathbf{K}_{0c}'|(1-K_L^2)} = \frac{2|\mathbf{K}_{sc}-\mathbf{K}_L||\mathbf{K}_{0c}-\mathbf{K}_L|\cos\theta}{|\mathbf{K}_{sc}-\mathbf{K}_{0c}|(1-K_L^2)}, \quad (45)$$

where

$$\cos\theta=R_L/Z_L. \quad (46)$$

Equations (41) and (45) can be used for determining the efficiency of four-terminal networks. When these equations are used, it is not necessary to know the absolute values of the angles for the complex reflection coefficients. It is sufficient to know their relative values. Also if it is more convenient to deal with impedances than with reflection coefficients, the impedances can be plotted on a Smith chart, or a $Z-\theta$ chart, and the magnitudes of the reflection coefficients used in Eqs. (41) and (45) can be found by making simple measurements on the chart.

JOURNAL OF APPLIED PHYSICS VOLUME 25, NUMBER 8 AUGUST, 1954

On the Chromatic Field Aberration of the Magnetic Electron Lens in the Electron Microscope

NOZOMU MORITO
Hitachi Central Research Laboratory, Tokyo, Japan
(Received November 9, 1953)

The chromatic field aberration of the magnetic lens in the electron microscope is theoretically studied. Good results are obtained for the coefficient of the chromatic difference in magnification of the objective, taking into account the inclination of the principal illuminating rays. The same coefficient of the intermediate lens increases rapidly as the magnification is reduced by the intermediate lens. The calculated results agree well with experimental results.

Calculated results are given for the coefficients of chromatic difference in rotation of all lenses as well as for the coefficient of chromatic difference in magnification of the projection lens.

INTRODUCTION

IN the electron microscope chromatic aberration, astigmatism, and spherical aberration are the most important defects of the image. The difference in color in light optics corresponds to the difference in electron speed in electron optics, and the inhomogeneity of the electron speed is caused mainly by the fluctuation of accelerating voltage and the energy loss of the electron in passing through a layer of matter. Since the chromatic aberration in the electron microscope has not been sufficiently compensated, great efforts have been made to stabilize the accelerating voltage. Fluctuation of the exciting current of a magnetic lens has effects similar to those of accelerating voltage, and the exciting current has also been kept extremely constant.

In our laboratory chromatic aberration has been studied for several years, and S. Katagiri has already reported his experimental results.[1] In this paper, the calculated results for chromatic field aberration of the magnetic lens will be described. The abstract has already been presented at the eighth lecture meeting of the Society of Electron Microscopy (Japan) held in Nagoya University on September 16, 1952.

CHROMATIC FIELD ABERRATION

The chromatic field aberration consisting of chromatic difference in magnification and chromatic difference in rotation causes the deviation of each image point from the Gaussian image point. The deviation is proportional to the relative variation of kinetic energy $\Delta\phi/\phi$ and the distance of the object point from the axis r_0. The radial component δ_{Fm} and the tangential component δ_{Fr} of the displacement referred to the object plane are repre-

[1] S. Katagiri, J. Electronmicroscopy **1**, 17 (1953).

sented as follows:

$$\delta_{Fm}=C_{Fm}\cdot r_0\cdot\frac{\Delta\phi}{\phi}, \tag{1}$$

$$\delta_{Fr}=C_{Fr}\cdot r_0\cdot\frac{\Delta\phi}{\phi}. \tag{2}$$

Here C_{Fm} and C_{Fr} are constants of proportionality known as the coefficient of chromatic difference in magnification and that of chromatic difference in rotation, respectively.

A variation in the exciting current Δi has the same effect as a change in kinetic energy of the electron $e\Delta\phi$ if

$$\frac{\Delta\phi}{\phi}=-2\frac{\Delta i}{i}. \tag{3}$$

CHROMATIC DIFFERENCE IN MAGNIFICATION

(1) For the Objective

If the specimen is illuminated by electron rays parallel to the axis, chromatic difference in magnification is proportional to the variation of the focal length of the lens $\Delta f_0/f_0$ due to an energy difference of the electrons $\Delta\phi/\phi$.

$$\left(\frac{\delta_{Fm}}{r_0}\right)_0=(C_{Fm})_0{}^{\mathrm{I}}\cdot\frac{\Delta\phi}{\phi}=-\frac{\Delta f_0}{f_0}. \tag{4}$$

Here, of course, it is assumed that the image distance is very large compared with the focal length. For the weak lens, the focal length is nearly proportional to ϕ and the coefficient is nearly equal to -1. However, as the lens strength increases, the value deviates from -1.

Assume that the axial field distribution of the lens is

$$H(z)=\frac{H(0)}{1+\left(\frac{z}{a}\right)^2}; \tag{5}$$

the focal length of the objective is given by

$$\left.\begin{aligned} f_0&=\frac{a_0}{\sin\frac{\pi}{\omega_0}},\\ \text{with}\quad \omega_0^2&=1+\frac{e}{8m\phi}H^2(0)a_0^2. \end{aligned}\right\} \tag{6}$$

Here e/m is the specific charge of electron and the suffix 0 represents quantities concerned with the objective. In the same manner, suffixes p and i will represent the quantities concerned with the projection lens and the intermediate lens, respectively. From Eqs. (4) and (6), the coefficient is given by

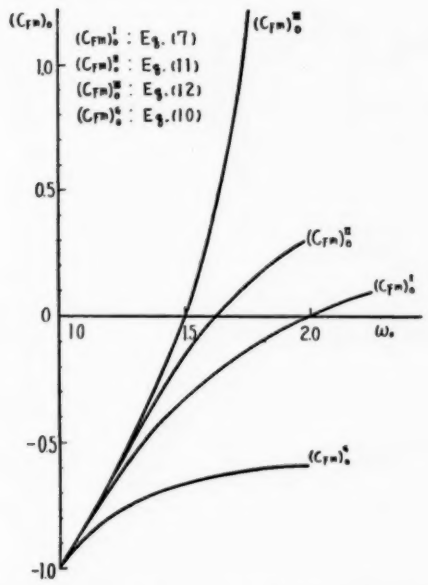


FIG. 1. Coefficient of chromatic difference in magnification for the objective (I).

$$(C_{Fm})_0{}^{\mathrm{I}}=\frac{\pi(\omega_0^2-1)\cos\frac{\pi}{\omega_0}}{2\omega_0^3\sin\frac{\pi}{\omega_0}}. \tag{7}$$

According to Eq. (5), Eq. (6), and the relation

$$4\pi ni=\int_{-\infty}^{\infty}H(z)dz, \tag{8}$$

the variable ω can be translated to the variable $ni/\sqrt{\phi}$ as follows.

$$\frac{ni}{\sqrt{\phi}}=16.81(\omega^2-1)^{\frac{1}{4}}\left[\frac{\text{amp. turn}}{(\text{volt})^{\frac{1}{2}}}\right]. \tag{9}$$

K. Kanaya[2] already reported Eq. (7) and W. Glaser[3] introduced a different formula:

$$(C_{Fm})_0{}^{G}=-\frac{\pi(\omega_0^2-1)}{2\omega_0^3\sin\frac{\pi}{\omega_0}}. \tag{10}$$

Figure 1 shows the relation between $(C_{Fm})_0$ and ω_0, and Fig. 2 the relation between $(C_{Fm})_0$ and $(ni/\sqrt{\phi})_0$. The experimental results show that the coefficient changes its sign at the value of $(ni/\sqrt{\phi})_0$ about 19 amp. turn/(volt)$^{\frac{1}{2}}$, while Glaser's coefficient $(C_{Fm})_0{}^{G}$ is

[2] K. Kanaya, Report of the Cooperative Research Committee on Electronmicroscopy of Japan, No. 56-A-7 (1951) (in Japanese).

[3] W. Glaser, Z. Physik. **117**, 285 (1941).

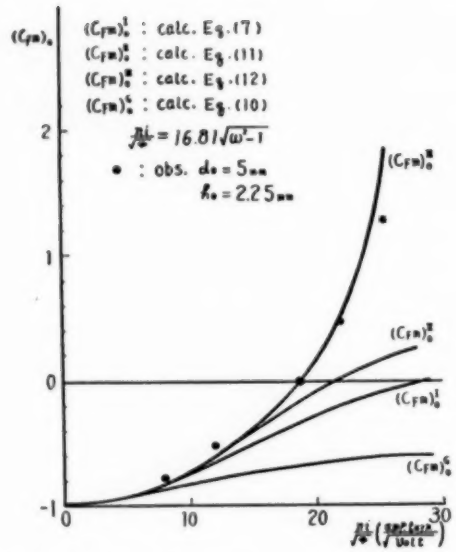


FIG. 2. Coefficient of chromatic difference in magnification for the objective (II).

always negative and $(C_{Fm})_0^{I}$ changes its sign at a very large value of $(ni/\sqrt{\phi})_0$. Thus, the results calculated from these formulas do not coincide with experimental results.

The disparity seems to be mainly caused by the assumption of parallel illumination. V. K. Zworykin and his co-workers[4] suggested that the inclination of the principal illuminating rays should be considered, but they scarcely described this quantitatively.

The author assumes that the inclination of a principal ray is equal to the inclination of a ray, at the object plane, which is parallel to the axis at a point sufficiently far from the lens.

If it is assumed that the deviation corresponding to $\Delta\phi$ of the plane conjugate to the image screen is Δf_0, it is possible to write

$$\left.\begin{aligned} &(C_{Fm})_0^{II} = (1+K_0)\cdot(C_{Fm})_0^{I}, \\ \text{with}\quad &K_0 = \frac{\omega_0 \sin\frac{\pi}{\omega_0}\cos\omega_0\pi + \sin\omega_0\pi\cos\frac{\pi}{\omega_0}}{\sin\omega_0\pi}. \end{aligned}\right\} \tag{11}$$

Here $(C_{Fm})_0^{I}$ is given by Eq. (7). The results of Eqs. (11) approach the experimental results as seen in Fig. 2.

It is more reasonable to use $(\Delta Z_f)_0$—the deviation of the focal point corresponding to $\Delta\phi$—instead of Δf_0. In this case, the coefficient is given by

$$(C_{Fm})_0^{III} = \left(1 - K_0 \sec\frac{\pi}{\omega_0}\right)\cdot(C_{Fm})_0^{I}. \tag{12}$$

The results from Eq. (12) fairly well coincide with the experimental results as seen in Fig. 2, where Eq. (9) is used for the translation of ω to $ni/\sqrt{\phi}$. The field distribution of the lens varies with the ratio, h/d, of the pole-piece spacing h to the pole-piece diameter d, and it is hence expected that the curve of the coefficient in Fig. 2 varies with h/d of the lens. This does not appear in the calculated results because of the assumption that the field distribution is given by Eq. (5). However, it is believed that a variation with h/d is practically out of the question, since lenses with a comparatively small value of h/d are usually used in the electron microscope.

(2) For the Projection Lens

If it is assumed that the chromatic difference in magnification depends only upon the variation of the focal length, the coefficient is expressed by

$$(C_{Fm})_p^{I} = \frac{\omega_p^2 - 1}{2\omega_p^2}\cdot(1 - \omega_p\pi\cot\omega_p\pi). \tag{13}$$

Figures 3 and 4 show the relations of $(C_{Fm})_p^{I}$ with ω_p and with $(ni/\sqrt{\phi})_p$, respectively. Contrary to the case of the objective, the calculated coefficient $(C_{Fm})_p^{I}$ agrees well with the observed values as seen in Fig. 4.

When the focal length of the projection lens is long, it is inadequate to consider the magnification as inversely proportional to the focal length. The correction is given by

$$\Delta(C_{Fm})_p = -\frac{1}{M_p}(C_{Fm})_p^{I}. \tag{14}$$

Here M_p is the magnification of the projection lens and its sign is chosen as positive for the erect image and negative for the inverted image. M_p is negative in general.

The principal planes of the projection lens separate

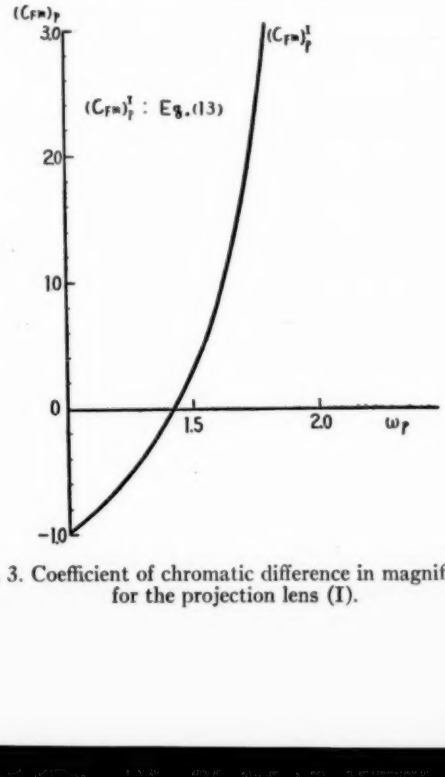


FIG. 3. Coefficient of chromatic difference in magnification for the projection lens (I).

[4] Zworykin, Morton, Ramberg, Hillier, and Vance, *Electron Optics and the Electron Microscope* (John Wiley & Sons, Inc., New York, 1945).

when the excitation is strong. A correction for the separation of principal planes is given by

$$\left.\begin{aligned} &\Delta(C_{Fm})_p = K_p' \cdot \frac{a_p}{l_{ps}}, \\ \text{with}\quad &K_p{}^1 = \frac{(\omega_p{}^2-1)(1+\cos\omega_p\pi)(\sin\omega_p\pi-\omega_p\pi)}{2\omega_p \sin^2\omega_p\pi}. \end{aligned}\right\} \tag{15}$$

Here l_{ps} represents the distance of the imaging screen from the projection lens. This correction need not be taken into consideration in practice, because a_p/l_{ps} is generally small and K_p' becomes large only when the excitation of the lens is stronger than for maximum magnification.

Correction for the inclination of the principal rays is considered similarly as for the objective. For the electron microscope with the lens system which comprises an objective and a projection lens, the correction is given approximately by

$$\left.\begin{aligned} &\Delta(C_{Fm})_p = K_p'' \cdot \frac{a_p}{l_{0p} - \dfrac{\omega_p a_p}{\tan\omega_p\pi}} \\ \text{with}\quad &K_p'' = \frac{(\omega_p{}^2-1)(\sin\omega_p\pi\cos\omega_p\pi-\omega_p\pi)}{2\omega_p \sin^2\omega_p\pi}, \end{aligned}\right\} \tag{16}$$

and for the electron microscope with three lenses, namely an objective, an intermediate lens, and a projection lens, it becomes approximately

$$\Delta(C_{Fm})_p = K_p'' \frac{\left\{1-\dfrac{(M_{0i})_m}{M_{0i}}\cdot\dfrac{l_{0i}}{l_{0p}}\right\}a_p}{l_{ip}-\dfrac{\omega_p a_p}{\tan\omega_p\pi}}. \tag{17}$$

l_{0p} and l_{ip} are the distance of the projection lens from the objective and from the intermediate lens, respectively, l_{0i} is the distance of the intermediate lens from the objective, M_{0i} is the magnification by the objective and the intermediate lens, and $(M_{0i})_m$ is the value of M_{0i} when the intermediate lens is not used. The value of $\omega_p a_p/\tan\omega_p\pi$ is usually small compared with l_{0p} and l_{ip}. In general, this correction is small and need not be considered.

For the projection lens, contrary to the objective, lenses with comparatively large h/d—the ratio of the pole-piece spacing and diameter—are available, as well as lenses with small h/d. Though Eq. (9) is utilized for the translation of ω_p to $(ni/\sqrt{\phi})_p$ in this study, it is not expected that the calculated results will agree with the experimental results for all lenses with various values of h/d. The fact that the value of $ni/\sqrt{\phi}$ at which the coefficient is equal to zero becomes small as the value of h/d becomes large has been proved by Katagiri's experiments in our laboratory. A similar tendency was found by the author in the relation between $ni/\sqrt{\phi}$ at minimum focal length and h/d. M. v. Ments and J. B. Le Poole[5] reported their fine work on the same subject. The result of the calculation, however, did not quantitatively coincide with that of Katagiri's experiment.

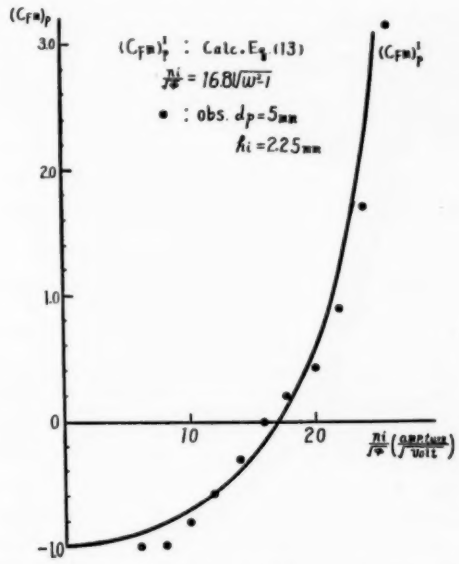


FIG. 4. Coefficient of chromatic difference in magnification for the projection lens (II).

(3) For the Intermediate Lens

The intermediate lens can be used not only similarly as the projection lens, but it can be used to reduce the magnification. It will be described here only for the latter case, since in the former case the intermediate lens is regarded as the first projection lens.

As the first approximation, it is assumed that the principal rays are parallel to the axis after passing through the intermediate lens. Under this assumption, the coefficient can be written as follows.

$$(C_{Fm})_i{}^{\mathrm{I}} = \frac{\Delta v_i}{v_i}\cdot\frac{\phi}{\Delta\phi} = \frac{v_i}{f_i}\cdot\frac{\Delta f_i}{f_i}\cdot\frac{\phi}{\Delta\phi}. \tag{18}$$

In general, the intermediate lens is used at a small value of $ni/\sqrt{\phi}$ and the variation of the focal length is nearly equal to that of kinetic energy of the electron:

$$\frac{\Delta f_i}{f_i} = \frac{\Delta\phi}{\phi}. \tag{19}$$

Consequently,

$$(C_{Fm})_i{}^{\mathrm{I}} = \frac{v_i}{f_i} = \left(1+\frac{v_i}{l_{0i}}\right)\left\{1-\frac{M_{0i}}{(M_{0i})_m}\right\} \tag{20}$$

[5] M. v. Ments and J. B. Le Poole, Appl. Sci. Research **B1**, 3 (1947).

is obtained, where v_i is the distance of the image plane from the lens.

Considering the inclination of the principal rays after passing through the intermediate lens, the corrected coefficient is given by

$$\left.\begin{aligned}(C_{Fm})_i^{II}&=\frac{1-R}{R}\\ \text{with}\qquad R&=\frac{M_{0i}}{(M_{0i})_m}.\end{aligned}\right\} \tag{21}$$

The coefficient is represented as a function of R and takes a simple and convenient form. R depends not only on the focal length of the intermediate lens, but on its position, so the coefficient cannot be written as a function of $ni/\sqrt{\phi}$ alone. If the lens is weak and it is permissible to write

$$\frac{1}{f_i}=C\left(\frac{ni}{\sqrt{\phi}}\right)_i^2, \tag{22}$$

the coefficient can be written in the form

$$\left.\begin{aligned}(C_{Fm})_i^{II}&=\frac{k\left(\dfrac{ni}{\sqrt{\phi}}\right)_i^2}{1-k\left(\dfrac{ni}{\sqrt{\phi}}\right)_i^2}\\ \text{with}\qquad k&=\frac{C\cdot l_{0i}\cdot v_i}{l_{0i}+v_i}\fallingdotseq\frac{C\cdot l_{0i}\cdot l_{ip}}{l_{0p}}.\end{aligned}\right\} \tag{23}$$

Figure 5 shows the variation of the coefficient with R

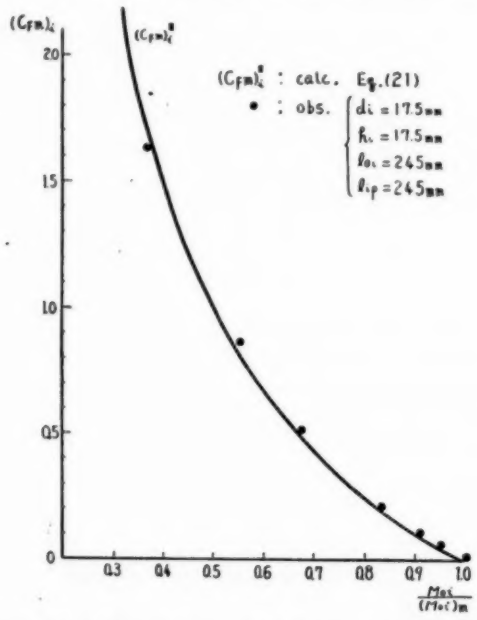


FIG. 5. Coefficient of chromatic difference in magnification for the intermediate lens (I).

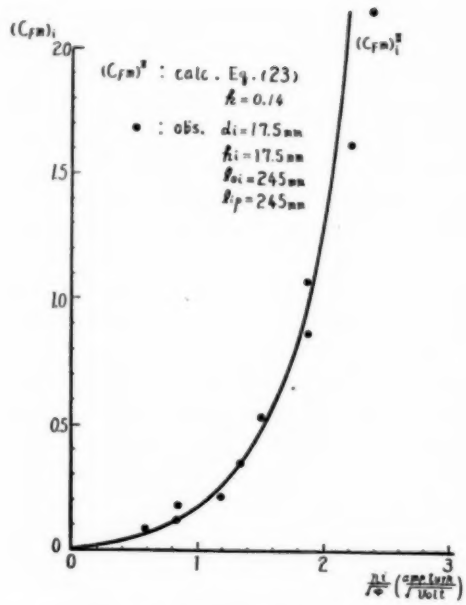


FIG. 6. Coefficient of chromatic difference in magnification for the intermediate lens (II).

and Fig. 6 the variation of the coefficient with $(ni/\sqrt{\phi})_i$ for a specific instrument.

For a strong lens, Eq. (19) is replaced by

$$\frac{\Delta f_i}{f_i}=p\frac{\Delta\phi}{\phi}, \tag{24}$$

and

$$(C_{Fm})_i^{III}=p(C_{Fm})_i^{II} \tag{25}$$

is obtained, where p is a function of ω_i; and if it is assumed that the field distribution of the lens is given by Eq. (5), it takes the form

$$p=-\frac{\omega_i^2-1}{2\omega_i^2}(1-\omega_i\pi\cot\omega_i\pi). \tag{26}$$

CHROMATIC DIFFERENCE IN ROTATION

(1) For the Objective

Image rotation due to the objective, when the field is assumed to have a form given by Eq. (5), is represented by

$$\psi_0=(\omega_0^2-1)^{\frac{1}{2}}\cdot\left(\cot^{-1}\frac{z_g}{a_0}-\cot^{-1}\frac{z_b}{a_0}\right), \tag{27}$$

where z_g and z_b are the z coordinates of the specimen plane and the correctly focused image plane, respectively. Differentiating by ϕ, under the assumption that the term inside the brackets is constant, leads to

$$(C_{Fr})_0=-\frac{\pi(\omega_0^2-1)^{\frac{1}{2}}}{2\omega_0}, \tag{28}$$

whereas Glaser[3] reported

$$(C_{Fr})_0{}^G=-\frac{\pi(\omega_0^2-1)^{\frac{1}{2}}}{2\omega_0{}^3}. \tag{29}$$

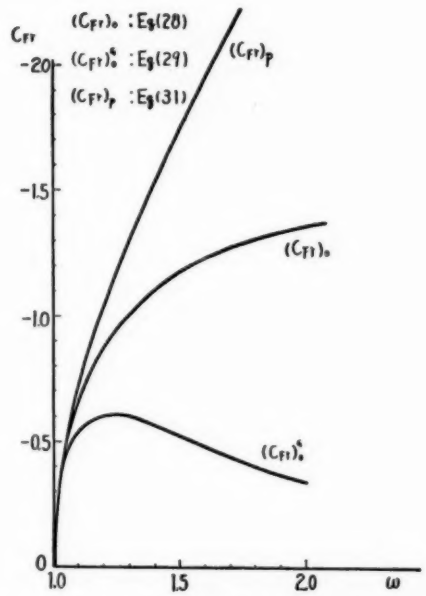


FIG. 7. Coefficient of chromatic difference in rotation (I).

In his calculation, the term inside the brackets is considered as a function of ω_0 or ϕ. Figure 7 shows the relation between $(C_{Fr})_0$ and ω_0 and between $(C_{Fr})_p$ and ω_p. In Fig. 8, the abscissa is translated to $ni/\sqrt{\phi}$ using Eq. (9).

Equation (28) is more suitable, in practice, than Glaser's formula (29), as is easily seen in the figure.

(2) For the Intermediate Lens and the Projection Lens

Chromatic difference in rotation of the intermediate lens is quite similar to that of the projection lens and, therefore, it will be described here for the latter alone.

Image rotation of the projection lens is given by

$$\psi_p=\pi(\omega_p{}^2-1)^{\frac{1}{2}}=\pi\left(\frac{2e}{m}\right)^{\frac{1}{2}}\cdot\left(\frac{ni}{\sqrt{\phi}}\right)_p. \qquad (30)$$

Differentiating by ϕ, as in the case of the objective, leads to

$$(C_{Fr})_p=-\frac{\pi(\omega_p{}^2-1)^{\frac{1}{2}}}{2}=-\frac{\pi}{2}\left(\frac{2e}{m}\right)^{\frac{1}{2}}\cdot\left(\frac{ni}{\sqrt{\phi}}\right)_p. \qquad (31)$$

Its behavior is shown in the above-mentioned Fig. 7 and Fig. 8. The calculated coefficient coincides approximately with the observed values. The coefficient $(C_{Fr})_p$ is linearly proportional to $(ni/\sqrt{\phi})_p$ and scarcely depends on the shape and dimension of the pole pieces.

(3) Sign of the Coefficient of the Chromatic Difference in Rotations

The coefficient of the chromatic difference in rotation calculated from Eq. (28) or (31) is always negative. A negative coefficient means that the tangential deviation of the image point due to an increase of the accelerating voltage is in the opposite direction to that of image rotation. Inversion of the excitation polarity causes a change of the direction of image rotation and, consequently, a change of the direction of the tangential deviation. It may be possible to include the sign in the representation of the coefficient, but no confusion will result in practice if the aforementioned facts are borne in mind.

COMPOUNDED ABERRATION OF THE LENS SYSTEM

The compounded chromatic field aberration of a lens system, which comprises an objective, an intermediate lens, and a projection lens, can be represented by use of the aberration coefficient of each lens.

For chromatic difference in magnification

$$\delta_{Fm}=\{(C_{Fm})_0+(C_{Fm})_i+(C_{Fm})_p\}\cdot r_0\cdot\frac{\Delta\phi}{\phi}, \qquad (32)$$

and for chromatic difference in rotation

$$\delta_{Fr}=\{(C_{Fr})_0\pm(C_{Fr})_i\pm(C_{Fr})_p\}\cdot r_0\cdot\frac{\Delta\phi}{\phi}. \qquad (33)$$

In Eq. (33), δ_{Fr} is positive when the direction of the tangential deviation of the image point is the same as the direction of the image rotation due to the objective. Positive and negative signs are chosen when the excitation polarity of the intermediate lens or the projection lens is the same and opposite to that of the objective, respectively. Of course, it is assumed that the optical axis of each lens is perfectly aligned.

Chromatic field aberration compounded of the chromatic difference in magnification and the chromatic difference in rotation is given by

$$|\delta_{Fmr}|=(\delta_{Fm}{}^2+\delta_{Fr}{}^2)^{\frac{1}{2}}. \qquad (34)$$

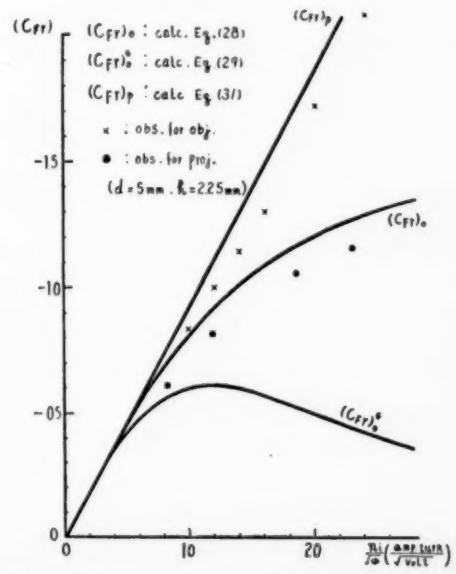


FIG. 8. Coefficient of chromatic difference in rotation (II).

With a knowledge of these relations and the behavior of each coefficient, problems concerning chromatic field aberration and its compensation in a lens system can be easily discussed.

DISCUSSION

Zworykin and his co-workers[6] describe the chromatic aberration as follows.

$$\Delta w_b=(C_1+iC_2)w_g+C_3w_B, \tag{35}$$

$$C_1=-M\cdot\frac{e\Delta\phi}{8m\phi^2}\int_{z_g}^{z_b}H^2r_\alpha r_\gamma dz, \tag{36}$$

$$C_2=-M\cdot\frac{\Delta\phi}{2\phi}\cdot\left(\frac{e}{8m\phi}\right)^{\frac{1}{2}}\int_{z_g}^{z_b}Hdz, \tag{37}$$

$$C_3=-M\cdot\frac{e\Delta\phi}{8mr_{\alpha B}\phi^2}\int_{z_g}^{z_b}H^2r_\alpha^2dz. \tag{38}$$

They deal with combined electric and magnetic fields, whereas only a pure magnetic field is covered here. Δw_b is the deviation of the image point from the Gaussian image point, w_g corresponds to the position of the object point and w_B corresponds to the position at which the ray passes the aperture plane, $z=z_B$. r_α and r_γ are the particular solutions of the paraxial ray equation

$$\ddot{r}+\frac{eH^2}{8m\phi}r=0 \tag{39}$$

and satisfy the boundary conditions

$$r_{\alpha g}=r_\alpha(z_g)=0,\quad \dot{r}_{\alpha g}=\dot{r}_\alpha(z_g)=1, \tag{40}$$

and

$$r_{\gamma g}=r_\gamma(z_g)=1,\quad r_{\gamma B}=r_\gamma(z_B)=0, \tag{41}$$

respectively.

Substituting Eq. (5),

$$\left.\begin{aligned}\cot^{-1}\frac{z_g}{a}&=\frac{\pi}{\omega},\\ \cot^{-1}\frac{z_b}{a}&=0,\\ \text{and}\quad \cot^{-1}\frac{z_B}{a}&=\frac{\pi}{2}\end{aligned}\right\} \tag{42}$$

in Eqs. (36), (37), and (38) leads to

$$C_1=-\frac{M\Delta\phi}{\phi}\cdot\frac{\pi(\omega^2-1)\cos\frac{\omega\pi}{2}}{2\omega^2\sin\frac{\omega\pi}{2}}, \tag{43}$$

$$C_2=-\frac{M\Delta\phi}{\phi}\cdot\frac{\pi(\omega^2-1)^{\frac{1}{2}}}{2\omega}, \tag{44}$$

$$C_3=-\frac{M\Delta\phi}{\phi}\cdot\frac{\pi(\omega^2-1)}{2\omega^2\sin\frac{\pi}{\omega}\sin\frac{\omega\pi}{2}}. \tag{45}$$

The first term in Eq. (35) depends only on the position of the object point and vanishes when the object point is situated on the axis, and the second term depends on the position at which the ray passes the aperture plane, namely on the ray path. C_1 is ascribed to chromatic difference of magnification, C_2 to chromatic difference in magnetic rotation, and C_3 to chromatic difference in image position.

These results are convenient in the case of dark field images in which the full free aperture is comparatively uniformly illuminated by the scattered electrons and the formation of image depends on the scattered electrons. In this case, the second term of Eq. (35) represents the disk of chromatic confusion and the first term represents the deviation of the center of the confusion disk from the Gaussian image point.

In many cases of the bright field image, however, scattered electrons form a very narrow pencil and illuminate only a part of the aperture. Electrons, which pass near by the object matter, also contribute to the formation of the outline of the image. In these cases, chromatic image defects observed in practice are different from those of the above-mentioned case. Namely, Δ_{Fm} given by Eq. (46) is observed as the chromatic difference in magnification.

$$\Delta_{Fm}=C_1w_g+C_3w_B^*. \tag{46}$$

Here w_B^* is the w_B of the principal ray.

If the principal ray is assumed to be a ray which is parallel to the axis at the point sufficiently far from the lens,

$$\left.\begin{aligned}r^*&=\frac{a\alpha\sin(\omega\varphi-\omega\pi)}{\sin\varphi}\\ \text{with}\quad \cot\varphi&=\frac{z}{a},\end{aligned}\right\} \tag{47}$$

[6] See reference 4, p. 570.

Δ_{Fm} is given by

$$\Delta_{Fm}=-\frac{M\Delta\phi}{\phi}\cdot r_g\cdot\frac{\pi(\omega^2-1)\cos\omega\pi}{2\omega^2\sin\omega\pi}$$

$$=\frac{M\Delta\phi}{\phi}\cdot r_g\cdot\left\{1-\frac{\omega\cos\omega\pi\sin\frac{\pi}{\omega}+\sin\omega\pi\cos\frac{\pi}{\omega}}{\sin\omega\pi\cos\frac{\pi}{\omega}}\right\}$$

$$\times\frac{\pi(\omega^2-1)\cos\frac{\pi}{\omega}}{2\omega^2\sin\frac{\pi}{\omega}}. \quad (48)$$

This result is quite the same as Eq. (12). If the aperture plane is chosen appropriately at the rear focus, Eq. (48) is derived directly from Eq. (36), but it must be considered that the position of the aperture plane varies with reference to the lens strength.

Chromatic aberration can also be solved by using a particular solution, r_β, instead of r_γ, which satisfies the boundary conditions[7]

$$r_\beta(z_g)=1, \quad \dot{r}_\beta(z_g)=0. \quad (49)$$

Chromatic aberration is then given, corresponding to Eqs. (35), (43), (44), and (45), by

$$\Delta w_b=(C_1'+iC_2')w_g+C_3'\dot{w}_g, \quad (50)$$

$$C_1'=\frac{M\Delta\phi}{\phi}\cdot\frac{\pi(\omega^2-1)\cos\frac{\pi}{\omega}}{2\omega^3\sin\frac{\pi}{\omega}}, \quad (51)$$

$$C_2'=-\frac{M\Delta\phi}{\phi}\cdot\frac{\pi(\omega^2-1)^{\frac{1}{2}}}{2\omega}=C_2, \quad (52)$$

$$C_3'=-\frac{M\Delta\phi}{\phi}\cdot\frac{a\pi(\omega^2-1)}{2\omega^2\sin\frac{\pi}{\omega}}, \quad (53)$$

where $\dot{w}_g$ represents the inclination of the ray at the object point. Here, Δ_{Fm} may be determined by

$$\Delta_{Fm}=C_1'w_g+C_3'\dot{w}_g^*. \quad (54)$$

Equation (48) can also be derived from Eq. (54), assuming the ray given by Eq. (47), since

$$\dot{w}_g^*=-\alpha\left(\omega\cos\omega\pi\sin\frac{\pi}{\omega}+\sin\omega\pi\cos\frac{\pi}{\omega}\right). \quad (55)$$

It may be convenient, for the case of the bright field image, to write the chromatic aberration as follows:

$$\Delta w_b=\left(C_1'+C_3'\frac{\dot{w}_g^*}{w_g}\right)w_g+C_3'(\dot{w}_g-\dot{w}_g^*). \quad (56)$$

The first term of Eq. (56) corresponds to the chromatic field aberration and the second term to the axial chromatic aberration, where $\dot{w}_g-\dot{w}_g^*$ is the inclination of the ray with respect to the principal ray.

The coefficient of axial chromatic aberration introduced by W. Glaser[3] corresponds to C_3'. His coefficient of chromatic difference in magnification has a different physical meaning because of the disregard of the fixed image plane.

CONCLUSION

For the chromatic difference in magnification of the objective, it is clear that the inclination of the principal illuminating ray should be taken into consideration. Though the assumed field—Eq. (5)—is not identical with practically realized distributions, it seems that the error of the results is not important. For the case in which the holes of the upper and lower pole pieces differ in diameter, it is necessary to calculate with an assumed field different from Eq. (5). The assumptions of the calculation for inclined principal rays appear to be reasonable for present-day electron microscopes.

For the projection lens, it is sufficient to consider that the chromatic difference in magnification is proportional to the variation of the focal length. In general, the correction mentioned in this paper need not be taken into consideration. The variation of the coefficient with the pole-piece diameter and spacing must now be inferred from experimental studies; it is to be hoped that a simple and practical method of calculating this variation will be found.

For the intermediate lens, the coefficient of the chromatic difference in magnification is always positive and increases rapidly from zero as the reduction ratio of magnification by the intermediate lens decreases from unity. In the three-stage electron microscope, complete correction of the chromatic difference in magnification cannot be attained unless the reduction ratio of magnification is greater than $\frac{1}{3}$. The permissible range of the ratio is more restricted when the objective and the projection lens are strong.

Granting that Glaser's coefficient of chromatic difference in rotation of the objective has some physical meaning, it cannot be used in practice. Experimental values smaller than those calculated may result from the magnetic potential drop in the iron yoke of the lens.

ACKNOWLEDGMENT

The author gratefully acknowledges his indebtedness to S. Katagiri and K. Koizumi, of Hitachi Central Laboratory, who kindly permitted him to quote their experimental results. He also wishes to thank Dr. B. Tadano, of the same laboratory, for helpful discussions and suggestions during the work.

[7] K. Kanaya, Bull. Electrotech. Lab. (Tokyo) **15**, 86 (1951) (in Japanese).

Statistical Hydrodynamics in Porous Media*

ADRIAN E. SCHEIDEGGER
Imperial Oil Limited, Calgary, Alberta, Canada
(Received December 2, 1953)

The statistics of disordered phenomena as exemplified by Einstein's theory of the Brownian motion is applied to the flow of fluids through porous media.

It is shown that such a statistical treatment of the hydrodynamics in porous media automatically explains some well-known phenomena in a more satisfactory manner than do capillaric models. The statistical theory leads to a differential equation of motion of the fluid which is a modification of that of Darcy; notably a new macroscopic quantity is introduced which is termed "dispersivity." This quantity is indicative of the sideways dispersion which a stream of fluid undergoes when it is passing through the porous medium. Under certain statistical assumptions outlined in the paper, the dispersivity becomes a constant of the porous medium.

The new differential equation of motion of the fluid is discussed in detail and some indications about applications are given.

I. INTRODUCTION

THE flow of fluids through porous matter is usually described by empirical laws of which the so-called law of Darcy in the form given by Muskat[1] is one of the best known. However, such empirical laws do not enable one to understand properly what is taking place in the porous medium, unless they can be founded upon more fundamental facts about flow through small openings.

It was realized long ago that Darcy's law must ultimately be deducible from the laws of viscous flow. Nevertheless, attempts of such a deduction have thus far been unsatisfactory. These previous attempts were mostly based (more or less openly) upon the construction of theoretical (capillaric) models: The porous medium is represented by a bundle of capillaries; the laws of viscous flow are applied to the latter and thus it is attempted to deduce relationships between the various macroscopically measurable quantities. A survey of a number of such theoretical models has been given and their implications have been discussed by the author[2] on an earlier occasion. The agreement of the properties of the capillaric models with those of natural porous material is rather poor; it is unavoidable to introduce fudge factors in order to get tolerable results. This, of course, is an unsatisfactory procedure and nullifies the very purpose of introducing the models.

A somewhat different approach to the problem was chosen by Kozeny.[3] In the latter's treatment, no open mention is made of any model. Nevertheless, all his considerations refer to cross sections normal to the direction of flow and pressure gradient in the porous medium—and are thus equivalent to the introduction of a model of parallel capillaries. Furthermore, recent experimental investigations[4] have shown that the relationships derived by Kozeny are quite inaccurate. In addition, results based on the models discussed above, and based especially on the Kozeny treatment, have been used to deduce fundamental relationships for the flow of fluids in porous media. This is particularly true for the attempts to correlate permeability and capillary pressure,[5] and relative permeability and capillary pressure.[6] These correlations were subject to experimental tests and showed rather large inaccuracies which, in turn, render the underlying Kozeny equation again more doubtful.

It therefore seems to the writer that an entirely different approach to the problem should be tried. Thus, instead of representing a porous medium by the extremely ordered device of a bundle of capillaries, it will be represented by a model of exactly opposite characteristics: A model that is assumed to be completely disordered. It is of course to be admitted that an actual porous medium will perhaps not be exactly represented by either type of models, but that it may lie somewhere in between the two. However, since the assumption of complete disorder appears never to have been rigorously investigated, it seems worth while to try it.

II. THE HYPOTHESIS OF COMPLETE DISORDER

The hypothesis of complete disorder of porous material will stand foremost in all further considerations. Naturally, it will not be mechanics in the ordinary sense that can be applied to such a disordered medium, but one will have to have recourse to statistical methods. The principal methods of the statistics of complete disorder have been initiated by Einstein in his paper on Brownian motion,[7] and later were expanded by

* Published by permission of Imperial Oil Ltd., Calgary.

[1] M. Muskat, *The Flow of Homogeneous Fluids Through Porous Media* (Edwards Brothers, Inc., Ann Arbor, Michigan, 1946).

[2] A. E. Scheidegger, Producers Monthly **17**, No. 10, 17 (1953).

[3] J. Kozeny, Ber. Akad. Wiss. (Wien) **136A**, 271 (1927).

[4] C. S. Brooks and W. R. Purcell, Trans. Am. Inst. Mining Met. Engrs. **195**, 289 (1952).

[5] W. R. Purcell, Trans. Am. Inst. Mining Met. Engrs. **186**, 32 (1949).

[6] L. A. Rapoport and W. J. Leas, Trans. Am. Inst. Mining Met. Engrs. **192**, 83 (1951).

[7] A. Einstein, Ann. Physik **17**, 549 (1905).

Langevin[8] and Smoluchowski.[9] A review of the methods is available in several papers.[10,11]

Statistics has been applied before in a very elementary way to flow in porous media.[12] In this instance, the process of passage of a fluid particle through a porous medium has been regarded as a stochastic process. However, this needs some explanation: In *one* particular piece of rock the path of an individual particle is most emphatically not the outcome of a stochastic process, but entirely determined by the prevailing boundary conditions and the equations of motion. In steady-state flow, each subsequent "particle" will follow *exactly* the path of the previous one, which is contrary to all stochastic assumptions. Thus, our fundamental idea is not to consider the statistics of flow in the first place, but to consider in lieu of one piece of a particular porous material, a great number (called "ensemble") of similar pieces with identical macroscopic properties and dimensions. In this way it is possible to define "averages over the ensemble" or "probabilities" for which dynamical equations may be set up. In each individual sample, however, the phenomena are governed by the ordinary laws of mechanics, i.e., they are entirely predetermined and the element of "chance" is excluded.

The fluid will be considered as a continuous medium, of which each point has a flow-path. We shall call each geometrical point of the fluid a "particle," but it should be understood that this has nothing to do with the actual molecules of the fluid. This is quite in contrast with the theory of Brownian motion where the statistics of the fluid and not those of the container are considered.

Thus, the fundamental problem in this paper will be to find the equations of motion for the probability density (over the ensemble) of the position of a particle. If one further assumption is made, which (for want of a better name) will be called "ergodic hypothesis," then it will be possible to identify the probability density with the actual physical density of a "parcel" of fluid represented by, say, a drop of dye spreading through the porous medium. This hypothesis assumes that the geometrical conditions for the motion of a particle prevailing at a spot in a particular porous medium are entirely uncorrelated with those at any other spot of that particular piece of material. This is actually an exact statement of what is meant by "hypothesis of disorder." Admittedly, no porous medium will behave in exactly that manner as the pore channels all have a finite length (and thus there is a correlation between geometries over such small pore lengths)—but this hypothesis is the only "backdoor" by which stochastics can be introduced into the present considerations. It seems to the writer that this hypothesis is at least as reasonable as assuming capillaric flow—although it corresponds to the opposite limit case.

Thus, we shall apply the statistical method as outlined before to the flow of fluids through porous media. Unless something else is specified, we shall make the following assumptions throughout the following sections of the paper:

(a) The porous medium is isotropic and, averaged over the ensemble, homogeneous.

(b) The external forces on the fluid are homogeneous and time independent. This may be realized by a constant pressure drop along a given direction. For the sake of simplicity, the gravity term will be neglected. It could easily be incorporated into the theory by introducing a more general external potential than the pressure alone.

(c) Different parts of *one* sample are "macroscopically identical." In other words, if a random selection of pieces is taken from only one sample, one obtains a series of samples forming an ensemble with which to do statistics. This means that a fluid particle proceeding through the porous medium finds the same ensemble probability for displacement at all points of its path. Under the assumptions (a) and (b) this also implies that the time average equals the ensemble average of the displacement of a particle. This is the assumption which has been called above "ergodic hypothesis." It is a rather restrictive hypothesis but it is held that it corresponds physically to a reasonable limit case.

In the final sections of this paper, hypotheses (a) and (b) will be relaxed; but this will be mentioned in the proper place. Without any further remarks, (a), (b), and (c) are understood to be assumed.

III. GENERAL THEORY

Consider a "particle" of fluid in the sense defined above. The probable positions of the particle at time t (where "probable" is meant with reference to the "ensemble of macroscopically identical porous media" and *not* with reference to the position of the particle in only one sample—the position of the particle in only one sample is sharply defined at all times by the Navier-Stokes equations) is now to be calculated.

We split the time 0 to t into N small intervals τ such that

$$N\tau = t. \tag{3.1}$$

In every time interval τ the particle will undergo a displacement ξ (with components ξ, η, ζ, with respect to a Cartesian coordinate system—vectors are denoted by boldface letters). The (ensemble-) probability that a given displacement ξ will take place in a time interval τ, will be denoted by $v(\xi)$. Under the assumptions made,

[8] P. Langevin, Compt. rend. **146**, 530 (1908).
[9] M. v. Smoluchowski, Ann. Physik **21**, 756 (1906).
[10] W. Pauli, *Statistische Mechanik* (Verein der Mathematiker und Physiker at Swiss Federal Inst. Technol., Zurich, 1947).
[11] L. Infeld, *On the Theory of Brownian Motion*, University of Toronto Studies, Applied Mathematics Series No. 4, Toronto (1940).
[12] E. C. Childs and N. Collis-George, Proc. Roy. Soc. (London) **A201**, 392 (1950).

this probability taken over the ensemble will depend neither on the position of the particle in the porous medium nor on the instant. The probability $v(\xi)$ is assumed to be normalized so that

$$\int v(\xi,\eta,\zeta)d\xi d\eta d\zeta = 1. \tag{3.2}$$

The average (over the ensemble) displacement in the time interval τ will be denoted by $\langle\xi\rangle_{Av}$. It is given by

$$\langle\xi\rangle_{Av} = \int \xi v(\xi) d^3\xi = (\bar{\xi},\bar{\eta},\bar{\zeta}). \tag{3.3}$$

It is very awkward to deal with probability distributions whose medians are not zero so that it will be convenient to introduce a coordinate system in which the median of the probability distribution will be zero. We shall indicate this coordinate system by dashes. Thus we set

$$\xi' = \xi - \langle\xi\rangle_{Av}. \tag{3.4}$$

In this dashed system the probability for the fluid particle to have a displacement ξ' will be

$$w(\xi').$$

The average of ξ' is zero. The function $w(\xi')$ may depend on $\langle\xi\rangle_{Av}$ as a parameter.

The fundamental question now is: What is the probability $w_N(x',y',z')$ over the ensemble in which a fluid particle with initial coordinates (0,0,0) in the dashed system will have coordinates x', y', z' after the time $t=N\tau$?

The problem thus appears as one of the random-walk type. It should be noted, however, that the problem is physically really very different from a random-walk phenomenon. We do *not* consider the path of only one particle in only one sample (that path is rigorously defined by the Navier-Stokes equations)—but rather the probable path of particles starting out of corresponding spots in an ensemble of porous media. As emphasized several times before, the statistics apply to porous media rather than to fluid flow.

Because of the assumptions that the porous medium is homogeneous and isotropic, and the field of forces homogeneous, the original probability density $w(\xi',\eta',\zeta')$ will be split into a product of three identical probabilities,

$$w(\xi',\eta',\zeta') = w(\xi')w(\eta')w(\zeta'), \tag{3.5}$$

and because of the ergodic hypothesis, those probabilities will be identical for every time step. In the light of the ergodic hypothesis, there are no limitations to the size of τ. This is, of course, somewhat of an oversimplification for very small τ, since the probability distributions w for 2τ and w for τ will be correlated. Nevertheless, assumption of the ergodic hypothesis greatly simplifies matters from a theoretical and physical standpoint, and describes porous media which are completely disordered over the smallest of intervals. It is therefore held as a reasonable limit case to allow τ to be arbitrary. Because of the ergodic hypothesis, and *only* because of this hypothesis, the problem of determining the probability of the position of the particle with respect to the ensemble becomes one of the random-walk type.

Thus, in terms of probability calculus, the problem is formally nothing but the composition of repeated trials. No matter what the original distribution w is, the central limit theorem of the calculus of probability[13] states that the distribution after N trials is Gaussian, provided N is large. Thus, if we denote by σ the standard deviation of w, it follows:

$$w_N(x',y',z') = (2\pi N\sigma^2)^{-\frac{3}{2}} \times \exp[-(x'^2+y'^2+z'^2)/2N\sigma^2)]. \tag{3.6}$$

The quantities σ and τ are constants during the motion (although σ may depend on $\langle\xi\rangle_{Av}$) so that one may set

$$\sigma^2/\tau = 2D, \tag{3.7}$$

where D is some factor which we shall call "factor of dispersion"—for a reason to be explained later. Using $N\tau=t$ yields,

$$w(t,x',y',z') = (4\pi Dt)^{-\frac{3}{2}} \times \exp\{-(x'^2+y'^2+z'^2)/(4Dt)\}, \tag{3.8}$$

and returning to undashed coordinates,

$$v(t,x,y,z) = (4\pi Dt)^{-\frac{3}{2}} \exp\{-[(x-\bar{x})^2 + (y-\bar{y})^2+(z-\bar{z})^2]/(4Dt)\}. \tag{3.9}$$

This is the fundamental probability distribution describing the journey of a fluid parcel through the porous medium. Of course, it is based upon all the statistical assumptions outlined in the previous sections.

IV. DARCY'S LAW

First we must find a connection between the average displacement $\langle\mathbf{x}\rangle_{Av}=(\bar{x},\bar{y},\bar{z})$ and the external field of forces. Because of the ergodic hypothesis one has

$$\langle\mathbf{x}\rangle_{Av}(t=N\tau) = N\langle\xi\rangle_{Av}. \tag{4.1}$$

Here, of course, use was made of the assumption that the external force was assumed to be homogeneous and time independent throughout the porous medium. According to these assumptions only, the average displacement is independent of the time and the location of the fluid parcel in the porous medium.

For each step τ, the laws of viscous flow state

$$(1/\tau)|\xi| = (B/\mu)|\mathrm{grad}p|\cos\theta, \tag{4.2}$$

if we denote the external force (per unit pore volume) by grad$p$ and the angle between ξ and $-$grad$p$ by θ;

[13] R. v. Mises, *Wahrscheinlichkeitsrechnung* (F. Deuticke, Wien, 1931).

B is a certain factor indicative of the reciprocal resistance of the opening through which the fluid (of viscosity μ) moves during the time step τ. The particular value of B at every point will depend on the shape and size of the opening through which the fluid packet passes.

The second problem is to form the average of (4.2) over "the ensemble." The average $\langle\xi\rangle_{Av}$ will be in the direction of $-\text{grad}p$ if the porous medium is assumed to be isotropic. Its magnitude will be equal to the average of the component of ξ in the direction of $-\text{grad}p$. The component of ξ in the $(-\text{grad}p)$ direction is equal to $|\xi|\cos\theta$. Thus, assuming that B and $\cos\theta$ are independent, we have,

$$\langle\xi\rangle_{Av}/\tau = -(\bar{B}/\mu)\,\text{grad}p\langle\cos^2\theta\rangle_{Av} \tag{4.3}$$

or

$$\langle\mathbf{V}\rangle_{Av} = -(\bar{B}/\mu)\,\text{grad}p\langle\cos^2\theta\rangle_{Av}, \tag{4.4}$$

where $\langle\mathbf{V}\rangle_{Av}$ is the intrinsic average velocity of the fluid; it is equal to $(1/P)$ times the macroscopic filter velocity $\mathbf{V}$ (P is the porosity). The average value of $\cos^2\theta$ (which is a constant of the material) may be called the reciprocal of the square of a "tortuosity," since $(1/\cos\theta)$ is only the relative length of the flow path per unit distance along the pressure gradient. We thus write

$$\langle\cos^2\theta\rangle_{Av} = 1/T^2. \tag{4.5}$$

From Eq. (4.3) the fundamental probability distribution becomes in virtue of (4.1)

$$v(t,x,y,z) = (4\pi Dt)^{-\frac{1}{2}}\exp[-(\mathbf{x}+t(\bar{B}/\mu) \times\text{grad}p\langle\cos^2\theta\rangle_{Av})^2/(4Dt)]. \tag{4.6}$$

The crest of this distribution proceeds with the velocity $\langle\mathbf{V}\rangle_{Av}$ which is the same as the intrinsic average velocity of the fluid. The filter velocity $\mathbf{V}=\langle\mathbf{V}\rangle_{Av}P$ is, therefore,

$$\mathbf{V} = -P(\bar{B}/\mu)\,\text{grad}p\langle\cos^2\theta\rangle_{Av}. \tag{4.7}$$

Comparing this with the so-called Darcy's law,[1]

$$\mathbf{V} = -(k/\mu)\,\text{grad}p, \tag{4.8}$$

(k=permeability, μ=viscosity) yields

$$k = P\bar{B}\langle\cos^2\theta\rangle_{Av}. \tag{4.9}$$

The $\langle\cos^2\theta\rangle_{Av}$ factor in the expression (4.9) for permeability has been introduced earlier by Sullivan and Hertel.[14] They arrived at it by considering the path of a particle through winding capillaries of only one sample of porous matter, and therefore their average is taken over the path of only one particle. Because of the ergodic hypothesis, this is the same as our ensemble average; the fact that the two methods agree serves as corroboration of the present formalism.

It is seen that a statistical theory gives one a connection between the macroscopic permeability and porosity, and the average of the microscopic reciprocal "resistance." Thus, one might call $\bar{B}$ the "intrinsic reciprocal resistivity" of the porous medium. Without the need to specify $\bar{B}$ further, the last formula shows that Darcy's law follows from a statistical theory for a porous medium, provided that proper assumptions for the statistics are made (ergodic hypothesis) and that the law of motion through each microscopic channel is such that the velocity of flow is proportional to the negative local pressure gradient and to the reciprocal viscosity of the fluid. Trivial as this statement may appear, it does not seem to have been proved before under such general conditions.

V. THE INTRINSIC RECIPROCAL RESISTIVITY

The quantity $\bar{B}$ cannot be specified further unless more specific assumptions about the porous medium are made.

If it be assumed that the flow channels are circular capillaries of diameter δ with the axis inclined at the angle θ towards the negative pressure gradient, the law of Hagen-Poiseuille gives for each individual capillary

$$B = \delta^2/32. \tag{5.1}$$

If this be averaged, one obtains

$$\bar{B} = \langle\delta^2\rangle_{Av}/32. \tag{5.2}$$

The expression for the permeability becomes

$$k = P\langle\delta^2\rangle_{Av}/(32T^2). \tag{5.3}$$

This is the expression which is also obtained by setting up a model of noninterconnected capillaries representing the porous matter.

One can rid oneself of the condition that the capillaries have to be circular. For a noncircular capillary of circumference γ the quantity B becomes

$$B = \gamma^2 C/(32\pi^2), \tag{5.4}$$

where C is a factor indicating the deviation of the cross section from a circle (C is 1 for a circle, 0 for a "fissure"). Thus one obtains for the reciprocal intrinsic resistivity

$$\bar{B} = (\bar{C}/32\pi^2)\langle\gamma^2\rangle_{Av}. \tag{5.5}$$

The average circumference of the channels can be expressed by the specific internal surface area S_Δ of the medium. If n be the number of pores per unit cross-sectional area normal to the pressure gradient, one has

$$\bar{\gamma}n/\langle\cos\theta\rangle_{Av} = S_\Delta, \tag{5.6}$$

but also

$$P = C'n(\pi/4)\langle\delta^2\rangle_{Av}/\langle\cos\theta\rangle_{Av} = C'n\langle\gamma^2\rangle_{Av}/(4\pi\cos\theta), \tag{5.7}$$

where C' is another deviation factor indicating the deviation of the pore cross section from a circle ($C'=1$ for a circle, 0 for a "fissure").

Thus, (interchanging $\langle\gamma^2\rangle_{Av}$ and $\langle\gamma\rangle_{Av}^2$)

$$\bar{\gamma} = \pi P/(C'S_\Delta), \tag{5.8}$$

[14] R. R. Sullivan and K. L. Hertel, *Advances in Colloid Science*. I (Interscience Publishers, Inc., New York, 1942), p. 37.

and

$$\bar{B}=(1/2)(\bar{C}/C'^2)(P^2/S_\Delta^2). \tag{5.9}$$

By comparison with Darcy's law,

$$k=(1/2T^2)(\bar{C}/C'^2)(P^3/S_\Delta^2). \tag{5.10}$$

The last equation is commonly termed "Kozeny equation."

The present deduction of the Kozeny equation from a theory that is applicable to a disordered porous medium shows just how many assumptions one has to introduce so as to arrive at it. Not only are "deviation factors" introduced, but also averages such as $\langle\gamma^2\rangle_{Av}$ and $\langle\gamma\rangle_{Av}^2$ are arbitrarily interchanged. The final undetermined factor $C/(2T^2C'^2)$ depends on properties of the porous matter and, through a proper microscopic arrangement, can have values in an extremely wide range.

One does not even need to take recourse to a statistical theory of the present type in order to show that the Kozeny equation is based on such arbitrary procedures as interchanging $\langle\gamma^2\rangle_{Av}$ and $\langle\gamma\rangle_{Av}^2$; a simple array of different-sized capillaries will bear this out.[12] The present deduction of the same fact from a statistical theory corroborates these earlier findings. It is, therefore, not at all surprising that the Kozeny relation gives such poor agreement with actual experimental tests.[12] As pointed out before, the Kozeny relation is equivalent to a capillaric model and more than a qualitative description of the empirical facts cannot therefore be expected from it.

VI. THE FACTOR OF DISPERSION

In the fundamental probability distribution there also appears the factor D. This factor does not enter into the laws of average advance of a fluid parcel in the porous medium; nevertheless, it has another macroscopic meaning.

The Gauss distribution at the time t is given by Eq. (4.6). It has been shown that the average position of the median of this distribution is

$$\langle\mathbf{x}(t)\rangle_{Av}=-\bar{B}\,\mathrm{grad}\,p\,t\langle\cos^2\theta\rangle_{Av}/\mu. \tag{6.1}$$

Similarly, the expectation value of the square length $\langle\mathbf{x}^2(t)\rangle_{Av}$ can be calculated. From the general theory of Gauss distributions it follows

$$\langle\mathbf{x}^2(t)\rangle_{Av}=2Dt+[(\bar{B}/\mu)\,\mathrm{grad}\,p\langle\cos^2\theta\rangle_{Av}t]^2. \tag{6.2}$$

This shows that the "dispersion" of a fluid particle within the ensemble is essentially dependent on D.

In virtue of the ergodic hypothesis, one can formulate these results so as to make them applicable to an extended fluid parcel within one sample of porous material. The proceeding of a fluid parcel through the porous medium could be realized experimentally: if a drop of dye were added to the liquid, one could watch the concentration of the dyestuff. Each particle of the drop of dye would proceed according to the probability distribution as given by Eq. (4.6). The concentration of dye at any one time is therefore given by the linear superposition of probability distributions of the type of (4.6), all starting out from different origins. The Eq. (4.6) is a solution of the differential equation,

$$\partial v/\partial t=\Delta(Dv)+\mathrm{div}[\bar{B}v\,\mathrm{grad}\,p\langle\cos^2\theta\rangle_{Av}/\mu], \tag{6.3}$$

as can be verified by a straightforward check by differentiation. The most general solutions of Eq. (6.3) is a linear superposition of solutions of the type (4.6), i.e., just the type of function which describes the concentration of the dye in the fluid. The fluid parcel, thus, proceeds according to Eq. (6.3).

In the present theory, since grad$p$ etc. have been assumed as constant, the factors of v could be pulled through the Laplacian and the divergence, and then Eq. (6.3) would look as follows,

$$\partial v/\partial t=D\Delta v+(\bar{B}\langle\cos^2\theta\rangle_{Av}/\mu)(\mathrm{grad}\,p\ \mathrm{grad}v). \tag{6.3a}$$

The statistical theory of flow through porous matter thus yields the result that each drop of liquid gets dispersed while it is passing through the medium. This is a fact which should be expected to occur, but which has never been described by any of the former theories of flow.

The factor of dispersion D was introduced in a somewhat formal way, as composed of the (ensemble-) standard deviation of the motion of a particle in an arbitrary time interval τ. It would be desirable to have another expression for D in terms of the microscopic quantities of the motion of each fluid particle.

This can be accomplished in the following way using a procedure established by Langevin.[8] Newton's equation of motion for a particle of the fluid is (force equals mass time acceleration),

$$B(\mathbf{f}-\rho\mathbf{x}^{\cdot\cdot})=\mathbf{x}^{\cdot}\mu. \tag{6.4}$$

This, of course, is a new basic assumption, but Newton's law of motion is generally accepted as valid. In (6.4) $\mathbf{f}$ is the force per unit volume (apart from the viscous forces), ρ the mass density and μ the viscosity of the fluid, B the local reciprocal resistance as defined before, and dots in juxtaposition denote time derivatives.

If one multiplies the last equation by $\mathbf{x}$, one obtains

$$B\mathbf{f}\mathbf{x}-B\rho\mathbf{x}^{\cdot\cdot}\,\mathbf{x}=\mathbf{x}\mathbf{x}^{\cdot}\mu. \tag{6.5}$$

Now one observes that

$$\mathbf{x}^{\cdot\cdot}\,\mathbf{x}=(d/dt)(\mathbf{x}\mathbf{x}^{\cdot})-\mathbf{x}^{\cdot 2}=(1/2)(d^2/dt^2)\mathbf{x}^2-\mathbf{x}^{\cdot 2}, \tag{6.6}$$

thus that Eq. (6.5) becomes

$$B\mathbf{f}\mathbf{x}-(1/2)B\rho(d^2/dt^2)\mathbf{x}^2+B\rho\mathbf{x}^{\cdot 2}=(\mu/2)(d/dt)\mathbf{x}^2. \tag{6.7}$$

The task is now to form the average of this expression (6.7) over "the ensemble." The forming of the average of $\mathbf{x}$, $\mathbf{x}^{\cdot}$, and $\mathbf{x}^{\cdot 2}$ is no difficulty if it be borne in mind that time derivation and averaging can be interchanged.

As to the average of $\mathbf{fx}$, the following argument can be made.

First, $\mathbf{f}$ and $\mathbf{x}$ are independent in an ensemble of isotropic and homogeneous porous media. Thus the term $\langle(\mathbf{fx})\rangle_{Av}$ will simply be equal to $\langle\mathbf{f}\rangle_{Av}\langle\mathbf{x}\rangle_{Av}$. Second, the (ensemble) average $\langle\mathbf{f}\rangle_{Av}$ of $\mathbf{f}$ can be calculated using the method of Langevin.[8] Thus the force $\mathbf{f}$ per unit volume on a particle at the point $\mathbf{x}$ can be split into a disordered and an ordered component. The disordered component will make the particles deviate from their average course and will cancel out on taking the average over the ensemble. The ordered force will be the same in all samples of the ensemble and will be parallel to the direction of $-\text{grad}p$. For a particle proceeding at an angle θ with the $-\text{grad}p$ direction, the force due to the pressure gradient (i.e., the *ordered* component) will be $-\text{grad}p\cos\theta$. The component parallel to the $-\text{grad}p$ direction of this is $-\text{grad}p\cos^2\theta$. If this be averaged over the ensemble, one obtains

$$\langle\mathbf{f}\rangle_{Av}=-\text{grad}p\langle\cos^2\theta\rangle_{Av}.$$

Thus, the average over the ensemble of Eq. (6.7) is

$$-\bar{B}\langle\mathbf{x}\rangle_{Av}\,\text{grad}p\langle\cos^2\theta\rangle_{Av}-(1/2)\bar{B}\rho(d^2/dl^2)\langle\mathbf{x}^2\rangle_{Av}+B\rho\langle\mathbf{x}'^2\rangle_{Av}=(\mu/2)(d/dt)\langle\mathbf{x}^2\rangle_{Av}. \quad (6.8)$$

From earlier results, we may insert $\langle\mathbf{x}(t)\rangle_{Av}=-(\bar{B}/\mu)\times\text{grad}p\langle\cos^2\theta\rangle_{Av}t$ (from Eq. (6.1)) and $|\mathbf{x}'|=-(B/\mu)\times|\text{grad}p|\cos\theta$ (from Eq. (4.2)); thus, the last equation becomes

$$[(\bar{B}/\mu)\,\text{grad}p\langle\cos^2\theta\rangle_{Av}]^2t-\tfrac{1}{2}(\bar{B}\rho/\mu)(d^2/dl^2)\langle\mathbf{x}^2(t)\rangle_{Av}+(\bar{B}\rho/\mu)\langle[(B/\mu)\,\text{grad}p\cos\theta]^2\rangle_{Av}=\tfrac{1}{2}(d/dt)\langle\mathbf{x}^2(t)\rangle_{Av}. \quad (6.9)$$

This is a differential equation for $\langle x^2(t)\rangle_{Av}$ whose stationary solution is [assuming $\langle x^2(t)\rangle_{Av}=0$ for $t=0$]

$$\langle x^2(t)\rangle_{Av}=2(\bar{B}\rho/\mu)[\langle[(B/\mu)\,\text{grad}p\cos\theta]^2\rangle_{Av}-[(\bar{B}/\mu)\,\text{grad}p\langle\cos^2\theta\rangle_{Av}]^2]t+[(B/\mu)\,\text{grad}p\langle\cos^2\theta\rangle_{Av}]^2t^2. \quad (6.10)$$

If this be compared with the earlier expression (6.2) for $\langle\mathbf{x}^2(t)\rangle_{Av}$, one obtains

$$D=(B\rho/\mu)[\langle((B/\mu)\,\text{grad}p\cos\theta)^2\rangle_{Av}-((\bar{B}/\mu)\,\text{grad}p\,\langle\cos^2\theta\rangle_{Av})^2]. \quad (6.11)$$

As B and $\cos\theta$ have been defined in microscopic terms, Eq. (6.11) gives the complete connection of the macroscopic factor D with microscopic quantities.

It is rather awkward to deal with D in an equation for v, as it is a function of the pressure gradient. However, the expression for D may be written as follows:

$$D=(\rho a/\mu^3)(\text{grad}p)^2, \quad (6.12)$$

with

$$a=\bar{B}[\langle(B\cos\theta)^2\rangle_{Av}-(\bar{B}\langle\cos^2\theta\rangle_{Av})^2]. \quad (6.13)$$

Here, a is a constant of the porous material, which in principle, could be measured macroscopically. Because of its connection with the dispersion of a parcel in the porous medium, we shall call this constant its "dispersivity."

It may be noted that a different interpretation of the first term in Eq. (6.11) can also be given. For, $(1/2)\rho\langle\mathbf{x}'^2\rangle_{Av}$ is some constant, equal to the average kinetic energy E per unit pore volume. Therefore, we have

$$(1/2)\rho\langle\mathbf{x}'^2\rangle_{Av}=E=(1/2)\rho\langle(B\,\text{grad}p\cos\theta)^2\rangle_{Av}/\mu^2, \quad (6.14)$$

and thus

$$[(B/\mu)\,\text{grad}p\cos\theta]^2=2E/\rho. \quad (6.15)$$

VII. INHOMOGENEOUS FORCES

The foregoing discussions pertain to a homogeneous field of external forces only. However, the case where the pressure is varying from place to place in the porous medium is of practical interest.

The transition from a homogeneous field to an inhomogeneous one is made along the usual lines: It is assumed that in the ϵ-neighborhood of a point, x, y, z the forces are "almost" homogeneous. In other words, there exists a neighborhood ϵ to each point x, y, z, within which the external forces change arbitrarily little. Within such a neighborhood all that has been said before applies.

It is thus clear that all the earlier equations apply which involve only the differentials of the displacements. Thus, the motion of fluid in an inhomogeneous field may be summarized as follows:

Consider a particle of fluid (as defined in Sec. 2) whose probability to be at time t at the location x, y, z, be denoted by $v(x,y,z,t)$. According to the ergodic hypothesis this is also the density of an actually larger parcel of fluid when it gets spread out through the porous matter.

The differential equation of motion for v is given by Eq. (6.3). Because of the variability of grad$p$, etc., the form (6.3) has to be used and not (6.3a). Substituting D from Eq. (6.12) and $\bar{B}$ from Eq. (4.9) yields

$$\partial v/\partial t=\Delta[(\rho a/\mu^3)(\text{grad}p)^2v]+\text{div}[v(k/\mu P)\,\text{grad}p]. \quad (7.1)$$

Here, a, k, and P are macroscopic constants of the porous medium (dispersivity, permeability and porosity, respectively) which may be defined in terms of microscopic quantities as outlined in the earlier sections of this paper. Contrariwise, ρ and μ are properties of the liquid (density and viscosity, respectively).

From the differential equation of motion it follows that a fluid particle in (x,y,z,t) moves with the average velocity $\langle\mathbf{V}\rangle_{Av}$, as can be seen from the solution 4.6:

$$\langle\mathbf{V}\rangle_{Av}=-(k/\mu P)\,\text{grad}p, \quad (7.2)$$

which corresponds to an average filter velocity $\mathbf{V}$ ($\mathbf{V}=\langle\mathbf{V}\rangle_{Av}P$),

$$\mathbf{V}=-(k/\mu)\,\text{grad}p. \quad (7.3)$$

However, this is no longer equal to the velocity of the center of mass of a large parcel of fluid moving through the porous medium, nor is it equal to the velocity of the average position of a fluid parcel. Those two quantities can no longer be calculated simply and generally. One has to solve Eq. (7.1) for each particular field of force and calculate the required averages.

Equation (7.1) is not able to give the complete solution of the problem of flow through porous matter, as it surmises that the pressure gradient be known at every point of the medium. There are 3 unknown functions (*viz.* v, ρ and p) and only one equation. Two more equations are therefore required to make the problem determined.

One of the two further equations which are required is immediately given by the (experimental) relationship between the pressure and the density of the fluid, *viz.*

$$\rho=\rho(p). \tag{7.4}$$

The last equation which is needed, will be furnished by the continuity condition.

Thus, the problem is to formulate the continuity condition in terms of the probability distribution function $v(\mathbf{x},t)$.

At time $t=t_0$ there will be a certain distribution of the fluid in the porous medium. Let this distribution be denoted by $w(\mathbf{x}_0,t_0)$ where the position coordinate is now denoted by $\mathbf{x}_0$. Naturally, $\mathbf{x}_0$ has the same range of values as $\mathbf{x}$. At time t (arbitrary), the "original" distribution of fluid $w(\mathbf{x}_0,t_0)$ will have the probability density $v(\mathbf{x},\mathbf{x}_0;t,t_0)$ to be at the spot $\mathbf{x}$ at time t. Here, $v(\mathbf{x},\mathbf{x}_0;t,t_0)$ is a "key" solution of the differential Eq. (7.1). It satisfies the boundary conditions and is a delta function for $t=t_0$ of the argument $\mathbf{x}-\mathbf{x}_0$. (The calculus applicable to delta functions has been discussed e.g. by Schwartz.[15]) It has therefore properties similar to a Green's function.

Thus, the distribution of fluid at time t will be given by $w(\mathbf{x},t)$ as follows:

$$w(\mathbf{x},t)=\int v(\mathbf{x},\mathbf{x}_0;t,t_0)w(\mathbf{x}_0,t_0)d^3x_0. \tag{7.5}$$

Now, the physical concept of continuity states that the distributions w must be proportional to the densities of the fluid times the local porosity and thus we can write Eq. (7.5) as follows:

$$P(\mathbf{x})\rho(\mathbf{x},t)=\int v(\mathbf{x},\mathbf{x}_0;t,t_0)\rho(\mathbf{x}_0,t_0)P(\mathbf{x}_0)d^3x_0. \tag{7.6}$$

This is the required continuity condition.

The continuity condition in the form written in (7.6) is rather awkward to deal with. One would like to have it in differential form instead of integral form.

This aim can be achieved by taking the time derivative with respect to t on both sides of Eq. (7.6), one thus obtains

$$P(\mathbf{x})\frac{\partial\rho(\mathbf{x},t)}{\partial t}=\int\frac{\partial v(\mathbf{x},\mathbf{x}_0;t,t_0)}{\partial t}\rho(\mathbf{x}_0,t_0)P(\mathbf{x}_0)d^3x_0. \tag{7.7}$$

However, the solution v is subject to the differential Eq. (7.1); one can thus insert $\partial v/\partial t$ from (7.1) into (7.7) to obtain

$$P(\mathbf{x})\frac{\partial\rho(\mathbf{x},t)}{\partial t}=\int\left\{\Delta\left[v\frac{\rho a}{\mu^3}(\operatorname{grad}p)^2\right]+\operatorname{div}\left[v\frac{k}{\mu P}\operatorname{grad}p\right]\right\}\rho(\mathbf{x}_0,t_0)P(\mathbf{x}_0)d^3x_0. \tag{7.8}$$

Upon letting $t=t_0$, one finally obtains

$$P\partial\rho/\partial t=\Delta[(P\rho^2a/\mu^3)(\operatorname{grad}p)^2]+\operatorname{div}[\rho(k/\mu)\operatorname{grad}p]. \tag{7.9}$$

The last step, of course, is based upon juggling around with the delta function where relationships such as

$$\phi\operatorname{grad}_x\delta(\mathbf{x}-\mathbf{x}_0)=-\phi\operatorname{grad}_{x_0}\delta(\mathbf{x}-\mathbf{x}_0)=\delta(\mathbf{x}-\mathbf{x}_0)\operatorname{grad}_{x_0}\phi \tag{7.10}$$

have been used.

In this manner, the problem of flow through porous media is solved by the three Eqs. (7.1), (7.4), and (7.9).

The theory of flow through porous media as it is outlined above could be modified in several ways. First, it is noted that the pressure is introduced as a given "external force"; no pressure fluctuations within "the ensemble" are therefore taken into account. This is not quite as general as one might desire, and thus a pressure probability distribution with corresponding fluctuations could be introduced instead of an external pressure. Second, the porous medium could be assumed as anisotropic. This would necessitate that the permeability and dispersivity of the porous medium be treated as tensors[16] instead of as scalars.

These modifications, however, are fairly straightforward and would add nothing to the actual physical ideas of the method of statistics of disorder. Introducing them would just drown the physical facts in a rather involved and lengthy scheme of mathematics.

VIII. CONCLUSIONS

The theory outlined above gives a description of the hydrodynamics in porous media. The method of the statistics of disorder yields the differential equation for the motion of a fluid through a porous medium; if certain factors are set equal to zero, the motion is the same as described by Darcy's law. However, if those factors are not set equal to zero, then a new macroscopic effect is described of which it is known that it must occur: the effect of dispersion. Individual par-

[15] L. Schwartz, Ann. télécommunications **3**, 135 (1948).

[16] J. Ferrandon, Génie Civil **125**, 24 (1948).

ticles of the fluid do not only move along the streamlines resulting from Darcy's law, but they also get dispersed sideways. This is expressed by a new quantity denoted by a and termed "dispersivity," which appears in the equations of motion. Under certain statistical assumptions outlined in this paper, the dispersivity is a constant of the porous medium. The fluid flow through porous media is therefore, in this instance, determined by three macroscopic constants of the latter (porosity, permeability, and dispersivity) and not by only two as assumed heretofore.

The introduction of the dispersivity "constant" was based upon a variety of assumptions. Some of these assumptions were statistical in nature, others mechanical. Of the statistical assumptions, the ergodic hypothesis as outlined in Sec. 2 is the most restrictive. Of the mechanical ones, the hypothesis that the fluid is a continuum and the flow always laminar merit being emphasized. Thus, molecular slippage or Brownian motion within the fluid are disregarded.

The treatment of dispersivity as a constant of the material is therefore conditional to these assumptions. Thus it may be expected that future experimental tests will show that the new flow equations (as is also the case for Darcy's law) are valid in certain pressure ranges only. Even for restricted pressure ranges, it might turn out that the assumptions made, especially the ergodic hypothesis, were too narrow. In that case, however, it is held that the present theory will be a description of what is physically valid in a limit case, with the latter being a reasonable approximation to physical reality.

The statistical method would also be applicable with fewer assumptions, but the end formulas would be much more involved than they already are. In order to show the workings of the proposed method, it has therefore been thought permissible to make those assumptions.

In conclusion, the writer wishes to acknowledge his indebtedness to Imperial Oil Limited, Calgary, for the permission to publish this study. Dr. John v. Neumann of the Institute for Advanced Study in Princeton, New Jersey, read and critically commented upon the work presented here; the writer wishes to thank him.

JOURNAL OF APPLIED PHYSICS VOLUME 25, NUMBER 8 AUGUST, 1954

Coincidence of Pulse Trains

HENRY D. FRIEDMAN
Electronics Laboratory, General Electric Company, Syracuse, New York
(Received November 12, 1953)

New methods for determining the time fraction of coincidence of separate periodic pulse trains are considered. The analysis applies solely to pulse trains whose fundamental periods are commensurable. For two pulse trains, the coincidence fraction is easily determined. For three or more pulse trains, a method for determining the largest possible error resulting from direct use of elementary probability theory is developed. Hence the question: "When does probability theory provide a close approximation to the coincidence fraction?" is to some degree answered.

I. INTRODUCTION

IN a paper[1] published in this journal, K. S. Miller and R. J. Schwarz stated and proved certain fundamental results in the theory of pulse train interference or coincidence. The purpose of this paper is to present further results in this theory. Some of these results make use of probability theory.

The coincidence problem might arise in communications—in the interference of signal pulse trains, for example. However, the theory is also applicable to a large variety of manufacturing problems, involving simultaneous operation of similar machines or conveyors of different periodicities. For instance, if one person is asked to oversee a number of machines, each of which periodically requires his attention, we might ask: How often will two or more machines simultaneously require attention?

Pulse train is abbreviated PT. By a *simple pulse train* (SPT) we shall mean a PT having exactly one pulse per period. Figure 1 illustrates two SPT's, of periods T_1 and T_2 and pulse widths p_1 and p_2, respectively, and an initial phase s. By a *general pulse train* (GPT) we shall mean a PT having one or more pulses per period.

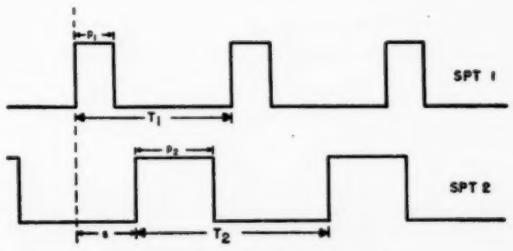


FIG. 1. Two simple pulse trains.

[1] K. S. Miller and R. J. Schwarz, J. Appl. Phys. **24**, 1032–1036 (1953).

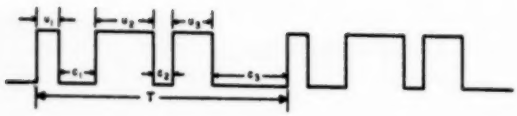


FIG. 2. A general pulse train.

(Every SPT is a GPT.) These pulses need not have equal width (see Fig. 2).

The work of the Miller-Schwarz paper is assumed for the demonstration of the following:

(a) A new method for solving the coincidence problem for two SPT's. This is a considerable simplification of the Miller-Schwarz method.

(b) The coincidence fraction for two GPT's can usually be approximated very simply by probability theory. An upper bound for the error is easily obtained.

(c) Using the above, the solution for three and four SPT's can be easily approximated. The method can be generalized to include five or more SPT's.

II. COINCIDENCE OF TWO SPT'S

Consider the two SPT's of Fig. 1. We shall assume, with Miller and Schwarz, that *our analysis applies solely to pulse trains whose fundamental periods are commensurable.* That is, we assume that all values are integral multiples of some number p:

$$p_1=\mu p \qquad T_1=mp$$
$$p_2=\nu p \qquad T_2=np$$
$$s=rp.$$

(As far as possible, the author has retained the Miller-Schwarz notation.)

The Miller-Schwarz method can be restated as follows: Let $g=gcd(m,n)$ and $M=lcm(m,n)$. We form all possible differences x_2-x_1, where

$$x_1=1, 2, \cdots, \mu$$
$$x_2=r+1, r+2, \cdots, r+\nu.$$

Let N be the number of these differences which are divisible by g. Then PT 1 and PT 2 will coincide the fraction

$$f=\frac{N}{M}=\frac{Ng}{Mg}=\frac{Ng}{mn}. \tag{1}$$

(The last equality results from the fact that $mn=gcd(m,n)\times lcm(m,n)$.)

The new method[2] follows, first for $r=0$ (zero lag at some point), then for the general case.

Let $r=0$. There exist integers a_1, a_2 such that

$$\mu=a_1g+r_1 \quad 0\leq r_1<g$$
$$\nu=a_2g+r_2 \quad 0\leq r_2<g. \tag{2}$$

[2] The referee has pointed out that the coincidence fraction of two pulse trains can be computed alternatively in a direct way from the cross-correlation function of the two trains. [For some general results, see D. Middleton, J. Appl. Phys. **22**, 1153 (1951), especially Sec. 4.]

Let $r_{max}=\max(r_1,r_2)$, $r_{min}=\min(r_1,r_2)$. Let f^k denote the coincidence ratio for $r=k$.

Theorem 1.—

$$f^0=\frac{\mu\nu+r_{min}(g-r_{max})}{mn}.$$

Proof.—We will count the number N of differences x_2-x_1 which are divisible by g. Table I sets a pattern which we can follow. Row by row, we associate with each x_2 all x_1 such that g divides x_2-x_1:

We count the differences as follows: For $x_2=1, 2, \cdots, g$ there are

$$d=r_1(a_1+1)+(g-r_1)a_1 \text{ differences} \tag{3}$$

which reduces to

$$d=r_1+ga_1=\mu. \tag{4}$$

The Table repeats itself every g rows. Hence for $x_2=1, 2, \cdots, a_2g$ there are $a_2\mu$ differences.

For $x_2=a_2g+1, a_2g+2, \cdots, a_2g+r_2$ there are

$$d'=\begin{cases} r_2(a_1+1)=r_2a_1+r_2 & \text{if} \quad r_2\leq r_1 \\ r_1(a_1+1)+(r_2-r_1)a_1=r_2a_1+r_1 & \text{if} \quad r_1\leq r_2 \end{cases} \tag{5}$$

or in either case

$$d'=r_2a_1+r_{min}. \tag{6}$$

Hence all values of x_2 give

$$N=a_2\mu+r_2a_1+r_{min}. \tag{7}$$

If we multiply both sides of (7) by g, and make use of (2), we obtain

$$f^0=\frac{Ng}{mn}=\frac{\mu\nu+r_{min}(g-r_{max})}{mn} \tag{8}$$

which proves the theorem.

If $r\neq 0$, we proceed as follows: Let PT 2* and PT 2** be defined as in Fig. 3. We can then apply Theorem 1 to PT 1 and PT 2*, and obtain

$$f^*=\frac{\mu(r+\nu)+r_{min}{}^*(g-r_{max}{}^*)}{mn}. \tag{9}$$

Similarly, we apply Theorem 1 to PT 1 and PT 2**, and

TABLE I.

x_2	x_1	Total per row
1	$1, g+1, 2g+1, \cdots, a_1g+1$	a_1+1
2	$2, g+2, 2g+2, \cdots, a_1g+2$	a_1+1
⋮	⋮	⋮
r_1	$r_1, g+r_1, 2g+r_1, \cdots, a_1g+r_1$	a_1+1
r_1+1	$r_1+1, g+r_1+1, \cdots, (a_1-1)g+r_1+1$	a_1
⋮	⋮	⋮
g	$g, 2g, 3g, \cdots, a_1g$	a_1
$g+1$	$1, g+1, 2g+1, \cdots, a_1g+1$	a_1+1
⋮	⋮	⋮
	etc.	

obtain

$$f^{**}=\frac{\mu r+r_{\min}{}^{**}(g-r_{\max}{}^{**})}{mn}. \tag{10}$$

It is easily seen that $f^r=f^*-f^{**}$ gives the coincidence ratio for the $r\neq 0$ problem. Hence we have proven

Theorem 2.—

$$f^r=\frac{\mu\nu+r_{\min}{}^{*}(g-r_{\max}{}^{*})-r_{\min}{}^{**}(g-r_{\max}{}^{**})}{mn}$$

where

$$\begin{aligned}
r_1{}^{*}&=r_1{}^{**}=\text{remainder of } \mu/g\\
r_2{}^{*}&=\text{remainder of } (r+\nu)/g\\
r_2{}^{**}&=\text{remainder of } r/g\\
r_{\min}{}^{*}&=\min(r_1{}^{*}, r_2{}^{*})\\
r_{\max}{}^{*}&=\max(r_1{}^{*}, r_2{}^{*})\\
r_{\min}{}^{**}&=\min(r_1{}^{**}, r_2{}^{**})\\
r_{\max}{}^{**}&=\max(r_1{}^{**}, r_2{}^{**}).
\end{aligned}$$

Corollaries.—Whenever $r=0$ and one (or both) of μ, ν is divisible by g, we have $r_{\min}=0$ and

$$f^0=\frac{\mu\nu}{mn}. \tag{11}$$

A similar conclusion may be drawn in the $r\neq 0$ case.

If $g=1$, then since all remainders must be less than g, we have $r_{\min}=r_{\min}{}^{*}=r_{\min}{}^{**}=0$ and

$$f^r=\frac{\mu\nu}{mn} \quad \textit{for all } r. \tag{12}$$

This result can be very useful. In a given pulse train experiment, it may be possible to choose period lengths which are prime to each other, and this would permit simple computations.

III. COINCIDENCE OF TWO SPT'S—THE PROBABILITY SOLUTION

The formula of Theorem 2 can be rewritten

$$\begin{aligned}
f^r&=\frac{\mu\nu}{mn}+\epsilon\\
\epsilon&=\frac{r_{\min}{}^{*}(g-r_{\max}{}^{*})-r_{\min}{}^{**}(g-r_{\max}{}^{**})}{mn}
\end{aligned} \tag{13}$$

$\mu\nu/mn$ (or p_1p_2/T_1T_2) is easy to obtain; we might call it the natural probability solution to the problem. Under what conditions is this a good solution?

Theorem 3.—For two SPT's

$$f^r=\frac{\mu\nu}{mn}\pm\epsilon \quad 0\leq\epsilon\leq g^2/4mn$$

where $g=gcd(m,n)$.

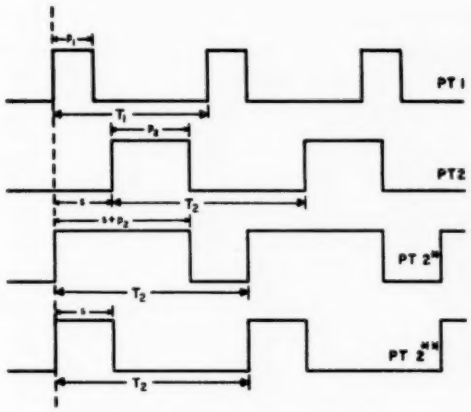


FIG. 3. Decomposition of PT 2.

Proof.—

$$0\leq r_{\min}(g-r_{\max})\leq r_{\max}(g-r_{\max}). \tag{14}$$

If g is even,

$$r_{\max}(g-r_{\max})\leq g/2(g-g/2)=g^2/4. \tag{15}$$

If g is odd,

$$r_{\max}(g-r_{\max})\leq\frac{g+1}{2}\left(g-\frac{g+1}{2}\right)=\frac{g^2-1}{4}. \tag{16}$$

In either case, we may say that

$$0\leq r_{\min}(g-r_{\max})\leq g^2/4. \tag{17}$$

Hence we have

$$f^r=\frac{\mu\nu}{mn}+(\epsilon^{*}-\epsilon^{**}) \qquad \begin{matrix}0\leq\epsilon^{*}\leq g^2/4mn\\ 0\leq\epsilon^{**}\leq g^2/4mn.\end{matrix} \tag{18}$$

The maximum error is not increased, but is merely allowed to be negative:

$$f^r=\frac{\mu\nu}{mn}\pm\epsilon \quad 0\leq\epsilon\leq g^2/4mn \quad QED. \tag{19}$$

We now have a condition under which the probability solution works: if $g\ll 2(\mu\nu)^{\frac{1}{2}}$, the approximation is usable. The main importance of the above solution, however, lies in its application to the work below.

IV. COINCIDENCE OF TWO GPT'S

We are given the general pulse trains GPT 1 and GPT 2, of periods T_1 and T_2, respectively. These trains have k_1 and k_2 pulses per period, of widths $u_1, u_2, \cdots, u_{k_1}$ and $v_1, v_2, \cdots, v_{k_2}$ (see Fig. 4). We assume that all pulse widths u_i and v_j, and all intervals c_i and d_j are integral multiples of some number p. (Again we require that all values bc commensurable.) Let

$$p_1=\sum_{i=1}^{k_1}u_i \quad p_2=\sum_{i=1}^{k_2}v_j \qquad \begin{matrix}T_1=mp\\ T_2=np.\end{matrix}$$

Then p_1 and p_2 are also multiples of p: $p_1=\mu p$, $p_2=\nu p$. Let f be the coincidence ratio for the two trains.

Theorem 4.—For two GPT's

$$f=\frac{\mu\nu}{mn}\pm\epsilon \qquad \epsilon\leq k_1k_2\frac{g^2}{4mn}.$$

Proof.—Let $u_i=\mu_i p$, $v_j=\nu_j p$, so that

$$p_1=\sum_{i=1}^{k_1}u_i=\left(\sum_{i=1}^{k_1}\mu_i\right)p=\mu p$$

or

$$\mu=\sum_{i=1}^{k_1}\mu_i,$$

and similarly

$$\nu=\sum_{j=1}^{k_2}\nu_j.$$

Consider the ith pulse of GPT 1 and the jth pulse of GPT 2. From Theorem 3, we know that the coincidence ratio for these two pulses is

$$\frac{\mu_i\nu_j}{mn}\pm\epsilon_{ij} \qquad 0\leq\epsilon_{ij}\leq\frac{g^2}{4mn}. \tag{20}$$

For fixed i, summation over j gives the coincidence ratio for the ith pulse of GPT 1 and all pulses of GPT 2. Summation of the latter over i then gives the desired coincidence ratio:

$$f=\sum_{i=1}^{k_1}\sum_{j=1}^{k_2}\left(\frac{\mu_i\nu_j}{mn}\pm\epsilon_{ij}\right)=\frac{\mu\nu}{mn}+\sum_{i=1}^{k_1}\sum_{j=1}^{k_2}(\pm\epsilon_{ij}). \tag{21}$$

But

$$\left|\sum_{i=1}^{k_1}\sum_{j=1}^{k_2}(\pm\epsilon_{ij})\right|\leq k_1k_2\frac{g^2}{4mn} \tag{22}$$

so that

$$f=\frac{\mu\nu}{mn}\pm\epsilon \qquad 0\leq\epsilon\leq k_1k_2\frac{g^2}{4mn} \qquad QED. \tag{23}$$

There is a more exact solution than the foregoing probability solution. First of all, when $g=1$ then (12) applies to GPT's also, and $\epsilon=0$. For $g\neq1$, it is possible to use Theorem 2 in a generalized form and thus obtain an *exact* answer. However, the computations will in general be tedious.

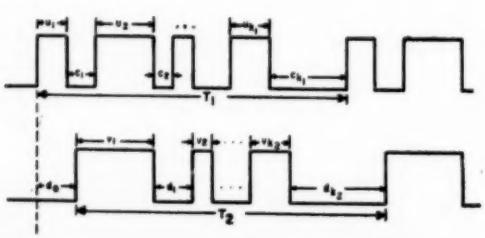


FIG. 4. Two general pulse trains.

V. THE PRODUCT OF TWO PULSE TRAINS

Any two trains GPT 1 and GPT 2 determine a third train GPT 12 in an obvious way. GPT 12 is the train whose pulses occur whenever GPT 1 and GPT 2 coincide (see Fig. 5). GPT 12 is called the *product* of GPT 1 and GPT 2.

Of special interest is the case where the two given trains are SPT's. We are given trains SPT 1 and SPT 2 (Fig. 1) and we form the product train GPT 12. GPT 12 will have k_{12} pulses per period, of widths $u_1, u_2, \cdots, u_{k_{12}}$. Let

$$p_{12}=\sum_{i=1}^{k_{12}}u_i \qquad \begin{matrix}T_1=mp\\ T_2=np\end{matrix} \qquad p_{12}=\mu_{12}p \qquad \begin{matrix}g=gcd(m,n)\\ M=lcm(m,n).\end{matrix}$$

Let f^r be the coincidence ratio for SPT 1 and SPT 2. (What follows is independent of initial phase relationships.)

Theorem 5.—$\mu_{12}=f^r\cdot M$.

Proof.—We know that

$$f^r=\frac{N}{M}. \tag{1}$$

It is easy to see that Np and $p_{12}=\mu_{12}p$ are both measures of the amount of coincidence of SPT 1 and SPT 2 in the common period Mp. (Mp is also the period of GPT 12.) Hence $Np=\mu_{12}p$ and

$$\mu_{12}=N=f^r\cdot M \qquad QED. \tag{24}$$

We will also need a value for k_{12}, the number of pulses per period of GPT 12. We shall obtain an estimate for this value, first when $r=s=0$ (or when SPT 1 and SPT 2 are in phase at some point).

Theorem 6.—Let $r=s=0$. Let a_1 and a_2 be defined for SPT 1 and SPT 2 as in (2). Then

$$a_1+a_2-1\leq k_{12}\leq a_1+a_2+1.$$

Proof.—Each pulse of GPT 12 is formed by the overlapping of pulses from SPT 1 and SPT 2. In every case, the left end of the overlap is determined by the left end of one, or both, of the SPT 1 and SPT 2 pulses (see Fig. 5).

In our integer model of two SPT's (Table I) this means that

$$k_{12}=n_1+n_2-1 \tag{25}$$

where n_i= the number of occurrences of $x_i=1$ in Table I. For Table I is essentially a description of all units of coincidence of SPT 1 with SPT 2. Each occurrence of $x_1=1$ can be interpreted as an overlap in which the left end of the SPT 1 pulse participates. Similarly for $x_2=1$ and SPT 2. We have counted $x_1=x_2=1$ twice; hence we subtract 1 from n_1+n_2.

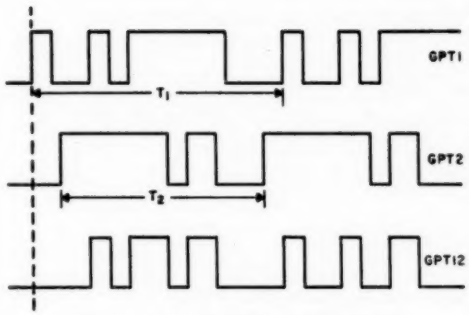


FIG. 5. GPT 12=GPT 1×GPT 2.

From Table I

$$n_2=\begin{cases}a_1+1 & 0<r_1\\ a_1 & 0=r_1\end{cases} \tag{26}$$

and by symmetry

$$n_1=\begin{cases}a_2+1 & 0<r_2\\ a_2 & 0=r_2.\end{cases} \tag{27}$$

Combining with (25) we have

$$a_1+a_2-1\leq k_{12}\leq a_1+a_2+1 \quad QED. \tag{28}$$

Corollary.—If $s\neq 0$ then $a_1+a_2-1\leq k_{12}\leq a_1+a_2+2$.

The proof is similar to the proof of the above theorem, except that (25) is replaced by

$$n_1+n_2-1\leq k_{12}\leq n_1+n_2. \tag{25'}$$

For if $s\neq 0$ we would have n_1=number of occurrences of $x_1=1$, n_2=number of occurrences of $x_2=r+1$. If g divides r, then there is a duplication as above, and we must subtract 1 from n_1+n_2; if g fails to divide r, there is no duplication, and $k_{12}=n_1+n_2$.

VI. COINCIDENCE OF 3 OR 4 SPT'S

We are now in a position to solve the coincidence problem for three or four SPT's. There is, first of all, an approximate solution with maximum error estimate. For the three-train problem, we proceed as follows: first we form the product of two of the trains. Using Theorem 6, we can find an upper bound for the number of pulses per period of the product train. By Theorem 5 (with Theorem 1 or 2) we can find the sum of the widths of all pulses of the product train. Finally, we use Theorem 4 on both the product train and the third SPT.

There are three ways to apply this procedure, depending on which pair of SPT's form the product train. There will be three final coincidence ratios and three error estimates, usually unequal. We need not compute all three answers, however, since the error estimates can easily be computed first, and the approach giving the least error chosen.

The four-train problem is approached in a similar manner. We have two pairs of SPT's, and we form the product of each pair. Theorem 4 is then applied to the two product trains. There will again be three approaches, and the least error approach can be chosen as before.

The maximum error estimate resulting from this method may be impractically large. When this happens, it is usually possible to arrive at an exact answer, but with a corresponding increase in computation. For the four-train problem, the exact method would proceed as follows: Using the Miller-Schwarz methods, we can obtain an exact description of the product train of two SPT's. That is, we can determine exactly where each pulse of the product train begins and ends. Hence, we can reduce the four SPT's to two completely described GPT's. The final comment of Sec. IV then applies: we can generalize Theorem 2 in exactly the same manner as Theorem 3 was generalized to Theorem 4. The answer thus obtained will be *exact*.

VII. COINCIDENCE OF 5 OR MORE SPT'S

The methods described above can be generalized to include the coincidence problem for any number of SPT's. It is desirable, for simplicity of computation, to use Theorem 4 (probability solution) as much as possible. Methods of solution can generally be designed so that pulse train pairs with large g occur in the earlier application of Theorems 1 and 2, and pairs having small g in the later probability application. In this way, the final error estimate is usually minimized. If the error is still too large, however, then exact methods must be used beyond the initial formation of produce trains. The remarks at the end of Sec. VI then apply.

The author is indebted to Mr. H. A. Samulon, Mr. K. Fong, and Mr. P. Weiss of the Electronics Laboratory, General Electric Company, for their advice and encouragement.

The Joint Distribution of n Successive Outputs of a Linear Detector*

WILLIAM C. HOFFMAN†
University of California, Los Angeles, California
(Received November 18, 1953)

The joint probability density function of n successive outputs of a linear detector is derived for the case of dependent Gaussian inputs.

1. INTRODUCTION

THAT a random noise input to a linear detector yields an output which is Rayleigh distributed is well known.[1,2] The independence of random noise from pulse to pulse permits us to write the joint probability density of n such outputs $R_1, R_2, \cdots, R_n$[3] as

$$f(r_1, \cdots, r_n)=\prod_{j=1}^{n}\frac{r_j}{\sigma^2}\exp\left(-\frac{r_j^2}{2\sigma^2}\right). \tag{1.1}$$

However, there exist cases, like "clutter" in a radar system, in which successive pulses are dependent, so that the joint probability density cannot be written in the relatively simple form (1.1). It is the purpose of this paper to derive, under the assumption of Gaussian inputs, the probability density function (hereafter abbreviated to p.d.f.) which applies in such cases.

The input to the linear detector is assumed to be the discrete-parameter Gaussian process (X_j, Y_j) $(j=\cdots, -1,0,1,\cdots)$. In electronics X_j and Y_j are termed "in-phase" and "in-quadrature" components, respectively, and may be thought of as real and imaginary parts of the complex Gaussian process $(Z_j)=(X_j+iY_j)$. The output is then of the form (under certain restrictions on the intermediate-frequency band width)

$$R_j=(X_j^2+Y_j^2)^{\frac{1}{2}}, \quad (j=\cdots,-1,0,1,\cdots) \tag{1.2}$$

and it is clear that (R_j) represents the modulus of the complex (Z_j) process.

2. DEFINITION AND PROPERTIES OF THE GAUSSIAN PROCESS

The (X_j,Y_j) process will be supposed stationary, i.e., the covariance function of (X_j,Y_j) and (X_k,Y_k) does not depend explicitly on j and k but only upon the difference $|j-k|$. It follows that the R_j process is also stationary in the wide sense.

The first- and second-order moments of the Gaussian process are then defined as follows:

$$EX_j=EY_j=0, \quad (j=\cdots,-1,0,1,\cdots) \tag{2.1}$$

$$EX_j^2=EY_j^2=\psi, \tag{2.2}$$

$$E(X_jX_k)=E(Y_jY_k)=\psi\rho_{|j-k|}, \quad j\neq k \quad (j,k=\cdots,-1,0,1,\cdots) \tag{2.3}$$

$$E(X_jY_k)=-E(X_kY_j)=\begin{cases}\psi\rho'_{|j-k|}, \\ 0, \quad j\neq k\end{cases}. \tag{2.4}$$

Let $\mathbf{\Lambda}_n$ denote the covariance matrix of the $2n$ random variables $X_1, Y_1, \cdots, X_n, Y_n$. According to the definitions (2.1)–(2.4), $\mathbf{\Lambda}_n$ must have the form

$$\mathbf{\Lambda}_n=\psi\mathbf{P}_n, \tag{2.5}$$

where $\mathbf{P}_n$ is the associated correlation matrix.

In the derivation of the joint distribution of $R_1,\cdots,R_n$, properties like (2.3) and (2.4) are required for the elements of the inverse correlation matrix $\mathbf{P}_n^{-1}$. The fact that these properties do hold is the content of the following lemma.

Lemma 2.1. Let the matrix $\mathbf{P}_n$ be nonsingular and defined by (2.3) and (2.4). Then the elements of $\mathbf{P}_n^{-1}$ are such that

$$\rho^{2j-1,2k-1}=\rho^{2j,2k} \tag{2.6}$$

$$\rho^{2j-1,2k}=-\rho^{2j,2k-1} \quad (j,k=1,2,\cdots,n) \tag{2.7}$$

$$\rho^{2j-1,2k}=-\rho^{2j,2k-1}=0 \text{ for } j=k. \tag{2.8}$$

Proof. It is clear from (2.3) and (2.4) that the following relations hold for the elements of $\mathbf{P}_n$:

$$\rho_{2j-1,2k-1}=\rho_{2j,2k}, \quad \rho_{2j-1,2k}=-\rho_{2j,2k-1} \quad (j,k=1,2,\cdots,n)$$

$$\rho_{2k-1,2k}=\rho_{2k,2k-1}=0 \quad (k=1,2,\cdots,n).$$

Furthermore, $\mathbf{P}_n$ is symmetric and nonsingular. Therefore, its complex isomorph,[4] $\mathbf{C}_n=(\rho_{2j-1,2k-1}+i\rho_{2j-1,2k})$ is Hermitian and non-singular. Consequently a unique $\mathbf{C}_n^{-1}$ exists and is also Hermitian, so that all its diagonal elements are real. By the same isomorphism as between $\mathbf{P}_n$ and $\mathbf{C}_n$ there is a one-to-one correspondence between

* A portion of a thesis submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy. It is a pleasure for the author to express his gratitude to Professor Paul G. Hoel for his aid and encouragement in this work.

† Author now at U. S. Navy Electronics Laboratory, San Diego 52, California.

[1] J. L. Lawson and G. E. Uhlenbeck, *Threshold Signals* (McGraw-Hill Book Company, Inc., New York, 1950).

[2] S. O. Rice, Bell System Tech. J. **23**, 282–332 (1944); **24**, 46–156 (1945).

[3] Throughout this paper capital letters will be used to denote random variables, and corresponding lower case letters will denote the associated real variables in the probability density function.

[4] Garrett Birkhoff and S. MacLane, *Modern Algebra* (Macmillan Company, New York, 1947), p. 217.

$\mathbf{C}_n^{-1}=(\rho^{2j-1,2k-1}+i\rho^{2j-1,2k})$ and the $2n\times 2n$ real matrix

$$\mathbf{P}_n^{-1}=\left[\begin{pmatrix} \rho^{2j-1,2k-1} & \rho^{2j-1,2k} \\ -\rho^{2j-1,2k} & \rho^{2j-1,2k-1} \end{pmatrix}\right].$$

Since products are preserved under isomorphism, it follows from $\mathbf{C}_n\mathbf{C}_n^{-1}=\mathbf{I}$ that $\mathbf{P}_n\mathbf{P}_n^{-1}=\mathbf{I}$. Inspection of the above form for $\mathbf{P}_n^{-1}$ then establishes the lemma.

3. THE JOINT DISTRIBUTION OF $R_1, R_2, \cdots, R_n$

The bivariate distribution of R_j, R_k is well known.[1,2] In the general n-dimensional case we start with the multivariate normal distribution of $X_1, Y_1, \cdots, X_n, Y_n$, which in view of Lemma 2.1 has the form

$$f(x_1,y_1,\cdots,x_n,y_n)=(2\pi)^{-n}\psi^{-n}|\mathbf{P}_n|^{-\frac{1}{2}}\exp\left\{-\frac{1}{2\psi}\sum_{j,k=1}^{n}[\rho^{2j-1,2k-1}(x_jx_k+y_jy_k)+\rho^{2j-1,2k}(x_jy_k-x_ky_j)]\right\}. \tag{3.1}$$

After the change of variables $x_j=r_j\cos\theta_j$, $y_j=r_j\sin\theta_j$ $(j=1,2,\cdots,n)$, the quadratic form in the exponent of (3.1) becomes

$$\sum_{j,k}[\rho^{2j-1,2k-1}r_jr_k\cos(\theta_k-\theta_j)+\rho^{2j-1,2k}r_jr_k\sin(\theta_k-\theta_j)]=\sum_{j=1}^{n}\beta^{jj}r_j^2+2\sum_{j<k}\beta^{jk}r_jr_k\cos(\theta_k-\theta_j-\varphi^{jk}),$$

where β^{jk} and φ^{jk} are defined as follows:

$$\begin{aligned}\rho^{2j-1,2k-1}&=\beta^{jk}\cos\varphi^{jk},\\ \rho^{2j-1,2k}&=\beta^{jk}\sin\varphi^{jk},\end{aligned}\qquad (j,k=1,2,\cdots,n).$$

So that $\beta^{jk}=\beta^{kj}$, $\varphi^{jk}=-\varphi^{kj}$, $\varphi^{jj}=0$. Consequently the desired marginal p.d.f. of $R_1, R_2, \cdots, R_n$ follows, in the form

$$f(r_1,\cdots,r_n)=(2\pi\psi)^{-n}|\mathbf{P}_n|^{-\frac{1}{2}}\prod_{j=1}^{n}r_j\exp\left(-\frac{1}{2\psi}\sum_{1}^{n}\beta^{jj}r_j^2\right)\cdot\int_0^{2\pi}\cdots\int_0^{2\pi}\exp\left[-\frac{1}{\psi}\sum_{j<k}\beta^{jk}r_jr_k\cos(\theta_k-\theta_j-\varphi^{jk})\right]\prod_{1}^{n}d\theta_j. \tag{3.2}$$

Now consider the multiple integral appearing in (3.2). Since the integrand is continuous, the multiple integral may be written as an n-fold iterated integral. Then induction on n, making successive applications of the formula[5]

$$e^{-z\cos\varphi}=\sum_{n=-\infty}^{\infty}(-1)^nI_n(z)e^{in\varphi},$$

to each iterated integral, and a final integration with respect to θ_n, permits the evaluation of the multiple integral. The p.d.f. of $R_1, R_2, \cdots, R_n$ then appears in the form

$$f(r_1,r_2,\cdots,r_n)=\psi^{-n}|\mathbf{P}_n|^{-\frac{1}{2}}\prod_{j=1}^{n}r_j\exp\left(-\frac{1}{2\psi}\sum_{1}^{n}\beta^{jj}r_j^2\right)$$

$$\times\sum_{\substack{\sum_{j<k}\nu_{jk}=\sum_{k<m}\nu_{km}\\ (k=1,2,\cdots,n)\\ 0\leqq\nu_{jk}\leqq\infty}}(-1)^{\sum_{j<k}\sum_{k=2}^{n}\nu_{jk}}\prod_{j=1}^{n-1}\prod_{k>j}\epsilon_{\nu_{jk}}I_{\nu_{jk}}\left(\frac{\beta^{jk}}{\psi}r_jr_k\right)\cos\left(\sum_{j=1}^{n-1}\sum_{k>j}\nu_{jk}\varphi^{jk}\right), \tag{3.3}$$

where ψ and $\mathbf{P}_n$ are defined by (2.2) and (2.5), respectively; $\epsilon_{\nu_{jk}}$ is the Neumann factor with value 1 for $\nu_{jk}=0$, and value 2 for $\nu_{jk}>0$; $I_\nu(x)$ is the modified Bessel function of the first kind with imaginary argument; and $\beta^{jk}=[(\rho^{2j-1,2k-1})^2+(\rho^{2j-1,2k})^2]^{\frac{1}{2}}$, $\varphi^{jk}=\arctan(\rho^{2j-1,2k}/\rho^{2j-1,2k-1})$, $(j,k=1,2,\cdots,n)$.

[5] G. N. Watson, *Bessel Functions* (Macmillan and Company, Ltd., London, 1944), rev. ed., p. 22.

Application of Ideal Gas Theory to the Gaseous Expansion from an Electric Spark*

R. B. EDMONSON, H. L. OLSEN, AND E. L. GAYHART
Applied Physics Laboratory, The Johns Hopkins University, Silver Spring, Maryland
(Received November 2, 1953)

The gaseous expansion from an electric spark is described in this paper in terms of ideal gas theory. The assumption is made that a quantity of energy is injected into a small gas volume in an infinitesimally short time and that the initial volume of gas then expands until it reaches the pre-spark pressure at an elevated temperature. Experimental results are given for a condenser discharge illustrating the dependence of the pressure equilibrated gas volume upon initial pressure and condenser energy. These results are interpreted using the ideal gas model presented. The significance of this expansion to the problem of spark ignition in combustible gases is discussed.

INTRODUCTION

THE problem of spark ignition in a combustible mixture is complicated by the fact that a multitude of physical and chemical processes is involved. The result is that the contribution of the electric spark cannot be evaluated fully since the heat energy of combustion is being added continuously at an unknown rate to the thermal energy transmitted initially to the gas by the spark. Information as to the role played by the spark energy alone in creating a local gas domain at a high temperature is obviously necessary before complete understanding of the more complex phenomenon of spark ignition can be achieved. Thus while the conditions necessary for ignition have been documented for several experimental procedures, there yet remains the problem of understanding the expansion and energy requirements of the hot gas kernel which follows the spark discharge and which in a flammable gas mixture precedes the establishment of a steady state combustion wave.

Rotating mirror measurements have been reported recently on the expanding channel of the spark by Craig and Craggs, Vul'fson and Libin, and Gegechkori. Craig and Craggs and Vul'fson and Libin[1] have studied the self-luminosity of the spark, while Gegechkori[2] has made measurements on schlieren photographs. These investigators have recorded the rate of expansion along the perpendicular to the direction of current flow for various values of spark circuit constants. Since the expanding gas possesses a rather strange symmetry, the volumetric expansion of the kernel cannot be determined from this type of measurement. Olsen, Edmonson, and Gayhart,[3] by photographing the kernel outline, have shown that the volumetric expansion approximates an adiabatic process from the time of shock separation to the time of pressure equilibration within the kernel. While it cannot be inferred from post-shock behavior that the observed pre-shock-separation expansion proceeds as an adiabatic process, there is little doubt that an appreciable fraction of the initial thermal energy of the spark is transferred from the kernel even though this process is not isentropic.

The aim of this paper is to describe the hot gas kernel in terms of its time-dependent volume and its energy content, utilizing ideal gas theory and the experimentally determined geometrical dimensions of the kernel. The objective of the experimental work is to ascertain whether the idealized theory can account for the kernel behavior to a first approximation under various conditions. A determination is made of the dependence of the pressure-equilibrated volume V_f upon condenser energy $1/2\,CX^2$ for an electrical system in which lead lengths external to the spark gap have been minimized and vacuum condensers have been used. In addition, the effect of the initial pressure upon the heated gas kernel has been determined for various gases.

IDEALIZED MODEL OF GAS KERNEL BEHAVIOR

In a consideration of the expansion of the spark kernel after shock wave separation, expressions have been derived for the energy content of the hot gas in terms of the initial gas conditions and certain measured volumetric quantities.[3] Using an adiabatic model for expansion, the internal energy of the kernel at the time of shock wave separation is given by Eq. (7) from reference 3

$$E_s=\left[\frac{P_0}{\gamma-1}\right]\left[\frac{V_f{}^{\gamma}}{V_s{}^{\gamma-1}}-V_0\right] \tag{1}$$

where the subscripts 0, s, and f refer, respectively, to the time of the spark, shock wave separation, and pressure equilibration. While the above equation refers specifically to the time of shock wave separation, it can be interpreted somewhat more generally in that the assumptions limit the development only to the time of thermal equilibrium. Thus, if at time (t) there is thermal equilibrium, this equation becomes

$$E_t=\frac{P_0}{(\gamma-1)}\left[\frac{V_f{}^{\gamma}}{V_t{}^{\gamma-1}}-V_0\right]. \tag{2}$$

* Work supported by Bureau of Ordnance, U. S. Navy, NOrd Contract 7386.

[1] R. N. Craig and J. D. Craggs, Proc. Phys. Soc. (London), **B66**, 500–511 (1953); K. S. Vul'fson and I. S. Libin, J. Exptl. Theoret. Phys. (USSR) **21**, 510–513 (1951). See J. M. Meek and J. D. Craggs, *Electrical Breakdown of Gases* (Oxford University Press, London, 1953), Chap. X for a complete discussion and bibliography.

[2] N. M. Gegechkori, J. Exptl. Theoret. Phys. (USSR) **21**, 493–506 (1951).

[3] Olsen, Edmonson, and Gayhart, J. Appl. Phys. **23**, 1157–1163 (1952).

If during the time for establishment of thermal equilibrium no appreciable expansion of the gases has occurred, then the thermal energy transferred from the spark to the gas can be written as

$$E_0=\frac{P_0}{(\gamma-1)V_0^{\gamma-1}}[V_f^{\gamma}-V_0^{\gamma}]. \tag{3}$$

Introducing a thermal efficiency coefficient defined by $E_0=\eta(1/2\,CX^2)$, then

$$\tfrac{1}{2}CX^2=\frac{P_0}{\eta(\gamma-1)V_0^{\gamma-1}}[V_f^{\gamma}-V_0^{\gamma}]. \tag{4}$$

From Eq. (2) the energy content of the kernel at the time of pressure equilibration is immediately obtainable [Eq. (8), reference 3] as

$$E_f=\frac{P_0}{(\gamma-1)}[V_f-V_0]. \tag{5}$$

The percentage of the initial thermal energy remaining in the ignition nucleus after the initial expansion depends upon the amount of thermal energy initially transferred and the initial volume of gas affected and can be obtained immediately. From Eqs. (3) and (5)

$$\frac{E_f}{E_0}=\frac{\beta-1}{\beta^{\gamma}-1}, \tag{6}$$

where β is the volumetric expansion ratio V_f/V_0. A plot of the ratio E_f/E_0 as a function of β is given in Fig. 1. Since the value of β decreases with a decrease in the initial heat content, the limiting value of the ratio E_f/E_0 as the initial energy is decreased is:

$$\lim_{E_0\to 0}\frac{E_f}{E_0}=\lim_{\beta\to 1}\frac{\beta-1}{\beta^{\gamma}-1}=\frac{1}{\gamma}. \tag{7}$$

A significant point is that there is an upper limit $(1/\gamma)$ to that fraction of the initial energy which remains within the ignition nucleus after this type of process and that limit is considerably less than unity.

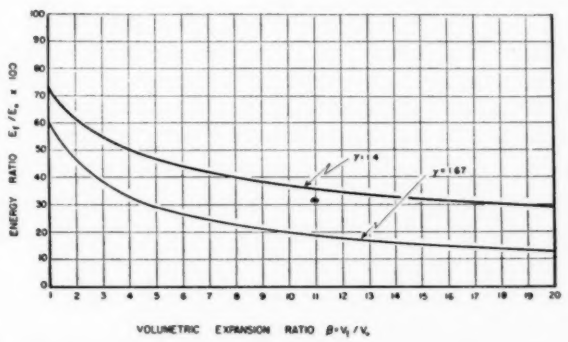


FIG. 1. Variation in the energy ratio as a function of volumetric expansion ratio for $\gamma=5/3$ and $7/5$.

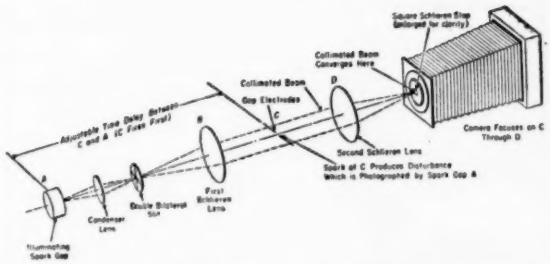


FIG. 2. Schematic diagram of the schlieren system showing the relative positions of the schlieren components.

EXPERIMENTAL

The primary data from which observations are made consist of time-controlled flash schlieren photographs of the ignition kernel subsequent to the time of pressure equilibration. Although the exact time interval from discharge to photographic exposure need not be known with great accuracy for this work, it is necessary that all schlieren photographs for a given series utilize the same delay interval so that individual photographs may be mutually and quantitatively compared. For the purpose of this experiment, twenty microseconds from spark to photograph is adopted as the standard time of observation.

Optical and Timing Arrangement

The schlieren system, Fig. 2, is a dioptric one utilizing 2-inch diameter, 8.5-inch focal length lenses. As a knife edge, a square opaque spot has been used together with a square aperture defining a light source, thereby producing radial symmetry in the schlieren image.[4] The illumination source is a Libessart spark gap with about 0.25×10^{-6} second effective duration. The time delay of twenty microseconds was introduced by an electronic interval timer triggered by electromagnetic radiation from the ignition spark gap.[5]

Gas Cell

The gas to be studied was placed in a closed cell which is shown by the line drawing and photograph of Fig. 3. Through the bottom of the cell passes an insulated metal rod into which is soldered one of the spark electrodes. This electrode is fixed in position with respect to the cell. Through the top of the cell passes the other (noninsulated) electrode holder. This holder has been provided with an expansion bellows and a differential screw which allows precise setting and alignment of the movable electrode with respect to the insulated one. At the ends of the cell are two glass windows with a diameter of 5 cm and a thickness of 6.4 mm, through which the schlieren beam is passed. At 90° to this beam axis, an observation window is available for electrode

[4] Edmonson, Gayhart, and Olsen, J. Opt. Soc. Am. **42**, 984 (1952).

[5] Edmonson, Gayhart, and Olsen, J. Soc. Phot. Engrs. **3**, 3, 135–144 (1952).

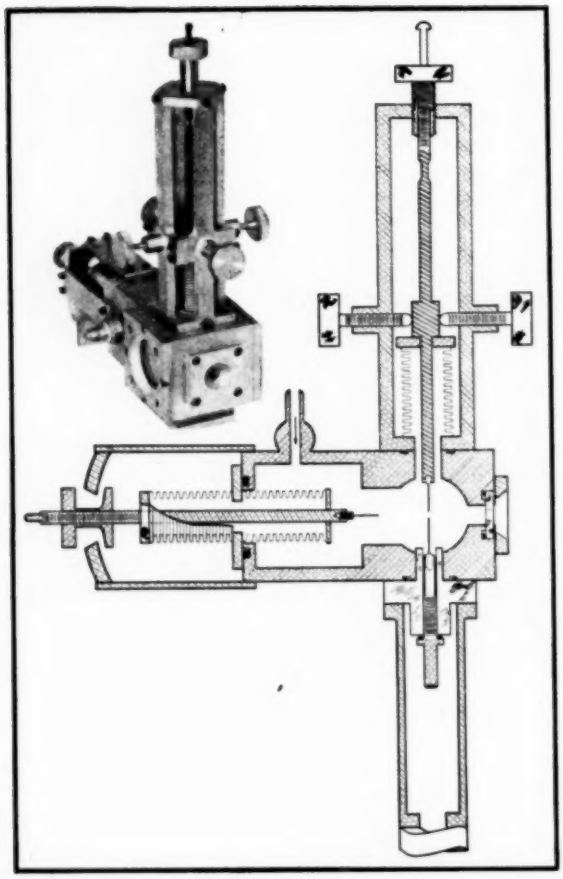


FIG. 3. Line drawing of the closed gas cell—inset—photograph of the cell completely assembled.

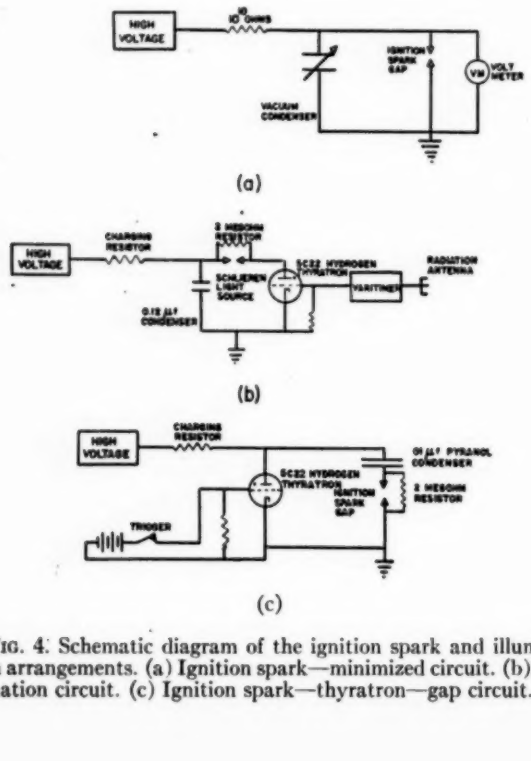


FIG. 4. Schematic diagram of the ignition spark and illumination arrangements. (a) Ignition spark—minimized circuit. (b) Illumination circuit. (c) Ignition spark—thyratron—gap circuit.

alignment purposes. The gap electrodes were aligned and the gap length measured to within 0.0001 inch using a Jones and Lamson optical comparator. On the side opposite the observation window is a double bellows arrangement which carries a rod of known diameter and an electrode stroking pin. The pin allows cleaning of the electrodes, a procedure which is occasionally necessary for the continued satisfactory operation of the spark. The rod provides a convenient means for determining the magnification of the schlieren system for each series of photographs, nominally 3.4. In some cases the stroking pin is replaced with a polonium ionization source† for stabilizing the gap breakdown voltage. The cell is mounted so that the axis of the electrodes, plated brass pins of 0.64-mm shank diameter, is vertical.

A gas supply, a mercury manometer, and a vacuum pump are connected to the gas cell. The argon, oxygen, and neon were from cylinders with a purity of 99.6 percent, 99.5 percent, and 99.9 percent, respectively. On introducing the test gas to the cell the gas chamber was repeatedly evacuated and pressurized until the dilution was negligible. The pressure was then adjusted to that desired for the particular experiment.

Electrical Methods

For the determination of the relationships between the condenser energy and equilibrated volume, a minimized circuit was employed. The term, minimized circuit, is used here to mean one for which the energy losses external to the spark gap have been reduced as much as is possible with the experimental arrangement. There is no control element in series with the ignition gap, and the electrical leads and connections are as energy loss free as the experimental arrangement will allow. Variable vacuum condensers were used for energy control inasmuch as the breakdown voltage for a particular gas should be determined only by the gas pressure and the gap length. This voltage was measured for each spark with a Sensitive Research electrostatic voltmeter. The capacity of the condenser was measured *in situ* with a General Radio capacity bridge, Type 716-B. The electrical arrangement is shown diagrammatically in Fig. 4(a). The gap length was arbitrarily set at 1.38 mm.

For the determination of the dependence of the spark kernel volume upon type of gas and the initial pressure the measurements at constant condenser energy could not be made under constant voltage, constant-capacity conditions in a minimized circuit in view of Paschen's Law for spark gap breakdown. However, there is no necessity for high efficiency of energy transfer for determination of this dependence. Since it was convenient, a hydrogen thyratron (Type 5C22) was used as a spark breakdown control element in series with the gap. This

† The polonium ionization source was kindly supplied by the Canadian Radium and Uranium Corporation, 630 Fifth Ave., New York, New York.

thyratron is relatively free of time jitter and is capable of passing high transient currents.[6,7] In this case, the ignition system was as shown in Fig. 4(c). The illumination arrangement, as shown in Fig. 4(b), remained the same for both experiments. Rather than initiation by the statistical breakdown of the ignition gap, the spark is initiated by "overvolting" the gap by triggering the hydrogen thyratron in that circuit. This spark discharge for the 0.01 μf pyranol condenser in the above arrangement has been found to have a duration of less than 0.5 microsecond.

RESULTS

The results of the experimental work may be classified into two general categories: those resulting from experiments performed in a minimized system to determine the effect of condenser energy upon the kernel volume and those resulting from experiments in the nonminimized system for the determination of the effect of the initial pressure upon the kernel volume.

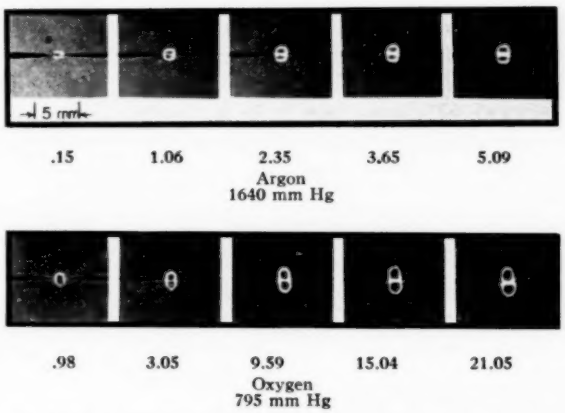


FIG. 5. Schlieren photographs of the pressure equilibrated gas kernel in argon and oxygen. Vacuum condensers—minimized circuit—twenty microseconds from spark to photograph—millijoules energy given under individual photographs.

The Effect of Condenser Energy on Kernel Volume

This series of measurements was carried out for argon and oxygen at a pressure of 795 mm mercury absolute and for argon at 1640 mm mercury absolute. Typical schlieren photographs are shown in Fig. 5. The kernels were reproducible and of regular outline. The results of the measurements are given in Figs. 6 and 7, in which the measured pressure-equilibrated volume has been plotted as ordinate against the vacuum condenser energy ($1/2\,CX^2$) as abscissa. In the case of argon, a supplementary polonium ionization source was used with the result that each argon series was taken under constant voltage conditions. The breakdown voltage did vary considerably from spark to spark for oxygen since supplementary ionization was not provided. The data suggest a power relationship between the condenser energy and the equilibrated volume as follows:

$$\text{For argon, both pressures} \quad 1/2\,CX^2 \propto V_f^{1.23}; \quad (8)$$

$$\text{For oxygen} \quad 1/2\,CX^2 \propto V_f^{1.57}. \quad (9)$$

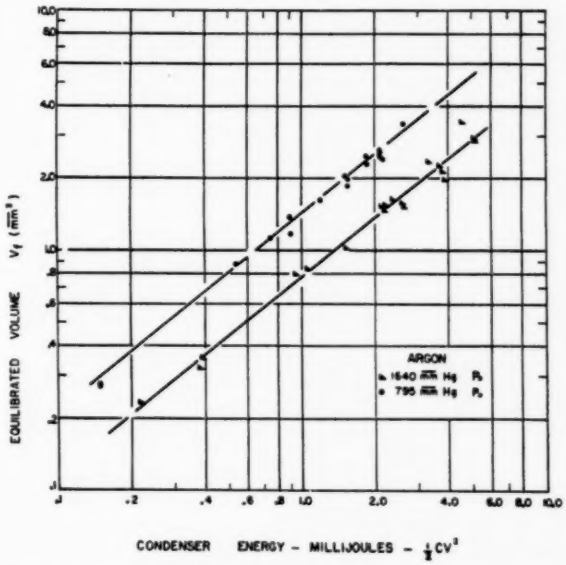


FIG. 6. Plot of equilibrated kernel volume as a function of condenser energy in a minimized circuit. Argon at initial pressure of 1640 and 795 mm Hg absolute.

The Effect of Initial Pressure on Kernel Volume

This series, for which the plotted data are given in Fig. 8, utilized neon, argon, and oxygen. The systematic

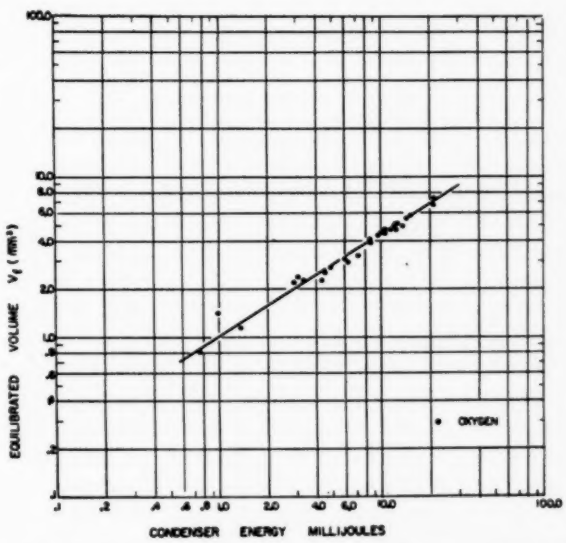


FIG. 7. Plot of equilibrated kernel volume as a function of condenser energy in a minimized circuit. Oxygen at an initial pressure of 795 mm Hg.

[6] W. T. Whelan, Trans. Am. Inst. Elec. Engrs. **67**, Part II, 1303–1307 (1948).

[7] K. J. Germeshausen, MIT Radiation Laboratory Series (McGraw-Hill Book Company, Inc., New York, 1948), Vol. 5, p. 335.

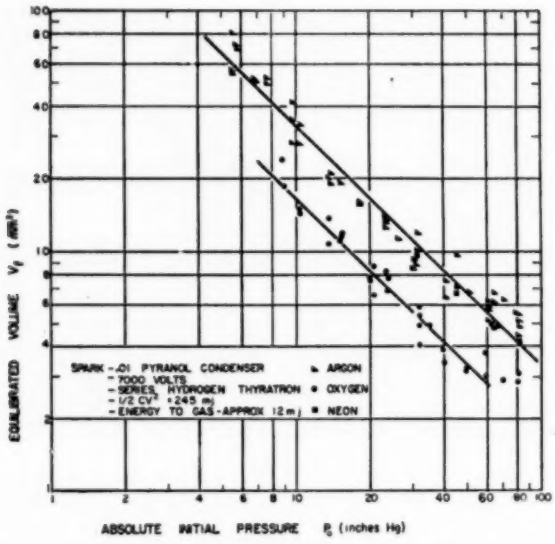


FIG. 8. Plot of equilibrated volume as a function of pressure. Constant energy—approximately 12 millijoules—delivered to spark gap.

dependence of the volume upon the initial pressure is shown. Straight lines with a negative slope of unity indicating the dependence $V_f \propto (1/P_0)$ have been inserted for each gas. These data, as previously mentioned, were not taken with a minimized system in that a 5C22 hydrogen thyratron was used in series with the gap and a pyranol condenser was used. However, the argon data in the minimized system for the two pressures do provide a kernel volume *vs* initial pressure check measurement. For this pressure ratio of 2.06, the inverse kernel volume ratio is 1.90 which is in conformity with the foregoing relationship. From a comparison of kernel sizes for the minimized and thyratron systems it can be seen that the latter system was producing an equilibrated kernel equivalent to that produced by a minimized system with a $1/2\,CX^2$ of approximately 12 mj.

DISCUSSION OF RESULTS

It is possible to make an interpretation of the experimental results for the monatomic gases which is consistent with ideal gas theory. Equation (4) expresses a theoretical relationship between the equilibrated kernel volume and the initial condenser energy for an efficiency transfer coefficient η. In this expression, it should be kept in mind that V_0, η, and γ all may, in general, vary in some manner depending upon the spark energy and the gas involved. Roth, Guest, von Elbe, and Lewis[8] have measured the energy release at a spark gap for monatomic gases for spark energies of the order of 1.0 millijoule in a minimized system and have concluded that it is approximately 0.95 $(1/2\,CX^2)$. While it does not follow necessarily that all of the energy dissipated at the gap is effective immediately in establishing a temperature in the gas kernel, for lack of specific information on this point the assumption is made here that the 95 percent energy transfer is valid. Using this value for η and assuming that γ does not vary appreciably for a monatomic gas, Eq. (4) can be solved for the initial volume V_0. With this result it now becomes necessary to consider whether the experimental results of Eq. (8) can be interpreted in a manner consistent with the theoretical Eq. (4). This equation may be written

$$\tfrac{1}{2}CX^2=\frac{P_0V_0}{\eta(\gamma-1)}\left[\left(\frac{V_f}{V_0}\right)^{\gamma}-1\right]. \tag{10}$$

If $[V_f/V_0]^{\gamma}$ is large as compared to one, then

$$\tfrac{1}{2}CX^2=\frac{P_0(V_f)^{(\gamma-1.23)}}{\eta(\gamma-1)V_0{}^{\gamma-1}}[V_f{}^{1.23}]. \tag{11}$$

Thus it is seen from the comparison with the experimental result [Eq. (8)] that if experiment and this theory are to describe the same physical phenomenon, then V_f and V_0 must vary in such a way that the coefficient of $V_f{}^{1.23}$ in Eq. (11) for a given pressure remains independent of energy. Thus the dependence of V_0 upon condenser energy is given by $V_0 \propto (CX^2)^{\frac{1}{3}}$.

Typical values for the initial volume V_0, the expansion ratio β, and the equilibrated energy-initial energy ratio E_f/E_0 are tabulated in Table I for an initial pressure of 795 mm Hg. Here, as computed from Eq. (6), the ratio E_f/E_0 varies according to the spark energy injected to the gas. In all cases, however, it should be noted that a major fraction of the initial energy has been utilized in the expansion function. These results are presented for monatomic gases only, for while oxygen satisfies the same type of power law, lack of knowledge of the dependence upon the condenser energy of the specific heat ratio γ and the fraction η of the electrical energy which appears as thermal energy instantaneously (as far as the expansion is concerned) makes it fruitless to follow through a parallel analysis for diatomic gases.

When $1/2\,CX^2$ is set as a constant of a system, it follows from Eq. (4) that the right member of this equation must be independent of the initial pressure of the gas. A sufficient condition that this be true is that

TABLE I. Typical values for initial volume, expansion ratio, and equilibrated energy-initial energy ratio at an initial pressure of 795 mm Hg.

Condenser energy	Equilibrated volume	Initial volume	Expansion ratio	Energy ratio
(mj)	V_f(mm³)	V_0(mm³)		E_f/E_0
0.1	0.215	0.042	5.12	0.289
0.2	0.375	0.061	6.15	0.263
0.4	0.67	0.094	7.13	0.242
1.0	1.43	0.160	8.94	0.213
2.0	2.52	0.235	10.72	0.191
4.0	4.45	0.348	12.79	0.171

[8] Roth, Guest, von Elbe, and Lewis, J. Chem. Phys. **19**, 1530–1535 (1951).

TABLE II.

Investigator	Gap width mm	Gas	Condenser type and range	Condenser energy	Vol. of gas Affected $\overline{mm}^3$ initial conditions
Coward Meiter	1.00	8.4% Methane-air at 752 $\overline{mm}$ Hg	Air 600 $\mu\mu f$	About 3.76 mj[a]	0.90
Edmonson Olsen Gayhart	1.38	Argon at 795 $\overline{mm}$ Hg	Vacuum 2123 $\mu\mu f$	2.6 mj	0.31
		Argon at 752 $\overline{mm}$ Hg		3.76 mj	0.36[b]

[a] Coward and Meiter do not give explicit values of voltage and capacity or condenser energy for their experiments. They do, however, quote Morgan's value of 0.0009 cal (3.76 millijoules) for the least igniting energy of a condenser spark in an 8.8 percent methane-air mixture and imply that their spark energy was little different from this.

[b] V_0 calculated by extrapolation of the 795 $\overline{mm}$ Hg pressure line of Fig. 6 and application of a pressure correction of 795/752.

each term be independent. Thus if in the subtractive term

$$\frac{P_0V_0}{\eta(\gamma-1)}$$

η and γ are not dependent upon pressure, $V_0 \propto (1/P_0)$. This result together with the term

$$\frac{P_0V_f^{\gamma}}{\eta(\gamma-1)V_0^{\gamma-1}}$$

predicts a dependence of $V_f \propto (1/P_0)$. This is the experimental result which has been obtained and is illustrated in Fig. 8 for both monatomic gases and oxygen.

Evidently independent data do not exist at this time which would serve as a check upon the reliability of the calculation of the initial volume of gas affected by the spark. One point of interest, however, is a comparison of the volume of argon initially affected by the spark with that volume of a methane-air mixture consumed by a sub-critical ignition spark as measured by Coward and Meiter.[9] Since there is some combustible consumed after the cessation of current flow in the spark, even though a sustained combustion wave is not established, it is to be expected that the total volume of gas affected will be greater than for a non-combustible gas. The experimental conditions and results for the two experiments are summarized in Table II. The two results are consistent and seem to be in reasonable agreement, particularly when the essential differences in the quantities measured are considered.

[9] H. F. Coward and E. G. Meiter, J. Am. Chem. Soc. **49**, 396–409 (1927).

CONCLUSIONS

Although the experimental work discussed has covered but a small number of the experimental variations possible, the various analyses, correlations, and computations with the experimental data have produced results which appear to be in conformity with the theoretical concepts based upon thermal equilibrium in an ideal gas. Therefore, assuming the sufficiency of application of the ideal gas theory to the processes in monatomic gases, the spark kernel energy can be derived from a knowledge of the pressure equilibrated volumes. For gases other than the monatomic ones, the lack of knowledge as to the specific heat ratio and dissociation at the temperature involved makes a similar discussion unjustifiable. However, the fact that the expansion process quickly removes an appreciable amount of energy from the spark kernel in a monatomic gas, projects the possibility that a parallelism exists in the ignition process. While the expansion may be considered as a fundamental process necessary for ignition of a gas by an electric spark, the energy of expansion, once lost from the gas kernel, cannot further participate in the mechanism of propagation of a combustion wave. It then follows that the electric spark, in effecting ignition in a flammable gas mixture, affects a localized mass of gas whose behavior is thereafter transient in character, and the energy transferred initially is employed in part usefully in performing its ignition function and in part uselessly in the regions of the gas mixture beyond that from which the flame originates.

ACKNOWLEDGMENTS

The authors acknowledge the helpful discussions of this paper with Dr. W. G. Berl, Dr. P. Rosen, Dr. R. Fristrom, Dr. A. Westenberg, and Mr. Rochelle Prescott. They also acknowledge the assistance of Mr. Fred Hammerschlag and Mrs. Frima Boyer in the design of some of the apparatus and of Mrs. Shirley St. Martin in the drawings and numerical computations.

Factors Determining the Permanent Magnet Properties of Single Crystals of Fe_2NiAl

E. A. NESBITT, H. J. WILLIAMS, AND R. M. BOZORTH
Bell Telephone Laboratories, Murray Hill, New Jersey

(Received February 17, 1954)

It is shown that fine precipitated particles present in single crystals of Fe_2NiAl are elongated and that the shape anisotropy of these particles is a major factor in determining the coercive force of Fe_2NiAl. For fields of less than 7000 oersteds, ⟨111⟩ directions are the directions of easy magnetization in the (110) plane of the crystal; this is due to the effect of the shape of the precipitated particles, which are oriented in three mutually perpendicular directions. In higher fields, the torque due to this effect decreases and then the [001] direction is the direction of easy magnetization because of intrinsic crystal anisotropy. These conclusions are based on an analysis, and on measurements, of the torque on three mutually perpendicular rods of magnetic material. These rods simulate the preciptated particles in the actual crystal.

THE most important permanent magnet alloys today (the Alnicos) are considered to be made up of fine particles[1,2] having dimensions of the order of 10^{-6} cm. The particles are too small to permit the existence of domain boundaries and therefore the processes of magnetization in these particles take place largely by rotation.[3] This is a high-energy process and it results in obtaining a high coercive force which may be calculated as

$$H_c \approx E/I_s,$$

where I_s is the saturation intensity of magnetization and E the energy associated with the anisotropy. In general, we may be dealing with various types of anisotropy that result from: the shape of the particle, the crystal energy, and the stress. Recently Néel[4] has proposed a new anisotropy depending on the surfaces of the particles.

The anisotropy of single cubic crystals of permanent magnet alloys has been measured by a number of investigators[5–7] using the torsion method. Briefly this consists of measuring the torque necessary to rotate a single-crystal disk specimen in a strong uniform magnetic field and then plotting the torque per unit volume as a function of the angular position of the crystal. From the measured values of torque, the magnetic anisotropy constant K is computed. For the (001) plane, the theoretical expression for the torque per unit volume is

$$L = -(K/2)\sin 4\theta,$$

where θ is the angle between the field and the [100] direction. For the (110) plane,

$$L = -(K/4)\sin 2\theta - (3K/8)\sin 4\theta.$$

The second anisotropy constant K_2 is here neglected.

In most cases for homogeneous single crystals these formulas are satisfactory for computing the anisotropy constants of the crystal. Permanent magnet crystals, however, have a heterogeneous structure[7] and therefore the situation is more complicated. This is illustrated by torque curves obtained on a single crystal* of Fe_2NiAl in the (100) plane as shown in Fig. 1. The torque is positive at an H of 5000 oersteds when the field angle θ is 25°, and it decreases with increasing H. At an H of 12 000 the torque is slightly negative and as H is further increased to a value of 25 000 the torque becomes increasingly negative. Torque curves for the (110) plane which act in a similar manner are shown in Fig. 2. This is an unusual behavior since the torque for homogeneous crystals, in general, rises with increasing H until magnetic saturation is reached and then remains constant.

These unusual torque curves, however, can be explained on the basis of the physical structure of the crystal. Figure 3 shows an electron micrograph of the (110) surface of a single crystal of Fe_2NiAl in which the material has a plate-like structure. It is this type of structure which is responsible for the unusual behavior of the torque curves shown in Figs. 1 and 2. It should be pointed out, however, that the structure shown in Fig. 3 was obtained by over-aging the crystal in order to make the plate-like structure visible in the electron microscope. The coercive force of this structure is 124 oersteds.

TORQUE CURVES ON WIRES OR RODS

The effect of a plate-like microstructure such as shown in Fig. 3 on torque curves was determined by studying analytically and experimentally, models simulating this microstructure. The analysis was simplified considerably by using elongated cylinders or wires instead of plates. Three models were made by imbedding a single wire, two wires crossed at 90°, and three mutually perpendicular wires in blocks of Lucite so that torque curves could be obtained on each structure. The wires were made of annealed molybdenum permalloy and had

[1] L. Néel, Compt. rend. **225**, 109 (1947).
[2] E. C. Stoner and E. P. Wohlfarth, Nature **160**, 650 (1947).
[3] C. Kittel, Phys. Rev. **70**, 965–971 (1946).
[4] L. Néel, Compt. rend. **237**, 23 (1953).
[5] N. V. Bugakov, Doklady Akad. Nauk S.S.S.R. **69**, 4 (1949).
[6] K. Hoselitz and M. McCaig, Proc. Phys. Soc. (London) **B64**, 549 (1951).
[7] R. D. Heidenreich and E. A. Nesbitt, J. Appl. Phys. **23**, 352 (1952); E. A. Nesbitt and R. D. Heidenreich, J. Appl. Phys. **23**, 366 (1952).

* Crystals were prepared as described in reference 7.

a length-to-diameter ratio of approximately 9. A torque curve on a single wire at $H=4000$ oersteds is shown by Curve 1 of Fig. 4. The curve is skewed and because of this skewness, if the curve is displaced 90° and replotted, the resultant of these two curves will not be zero. A curve of twice the frequency is then obtained for the resultant. The peak value of torque at this field strength is almost equal to that for the single wire but on a torque per unit volume basis it is reduced by a factor of two. Figure 5 shows the torque curve obtained on a single wire at $H=7000$ oersteds. Note that the curve is now skewed less at this higher field strength and that the resultant curve for 2 wires is consequently of lower amplitude. If H is increased further, the torque curve for a single wire tends to become symmetrical in the 90° interval and the resultant of the two curves, which are displaced 90° from each other, approaches zero. From this it is apparent that the torques on two single wires displaced 90° from each other tend to become equal in magnitude and since they are of opposite sign, the resultant torque approaches zero.

Torque curves (Fig. 6) taken directly on crossed wires that are rotated about an axis perpendicular to the plane in which the wires lay agree, of course, with the resultant of the two single wires displaced 90° from each other. It should be noted that the easy direction of magnetization for these crossed wires occurs when the direction of the field bisects the angle between the wires. This corresponds to a [110] direction in the (001) plane in the case of the crystal. The amplitude of the torque curves decreases and the skewed appearance of the curves tends to disappear with increasing field. It is shown analytically in Appendix I that the torque due to wires crossed at 90° approaches zero as the field becomes very large.

For the length to diameter ratios of the plate-like structures that may be encountered in permanent magnet crystal structures, the H required to obtain zero torque is very high. This is indicated by the high fields ($H=15\,000$) used to obtain the data in Fig. 6, and even these fields were not sufficient to saturate the wires. In addition, the crystal (Fig. 2) was not saturated for values as high as 31 000 oersteds. Further insight into the problem may be obtained analytically as shown in Appendix II. The curves for crossed wires obtained analytically are shown in Fig. 7 and they agree closely with the experimental curves. It is evident from the curves and from Eq. (6) in the Appendix that an infinite value of field would be required to reduce the torque of these crossed wires to zero.

Torque curves obtained on the model having three mutually perpendicular wires are shown in Fig. 8. The model was oriented so that the axis of rotation was perpendicular to a plane corresponding to a (110) plane in the crystal. These torque curves have the same sym-

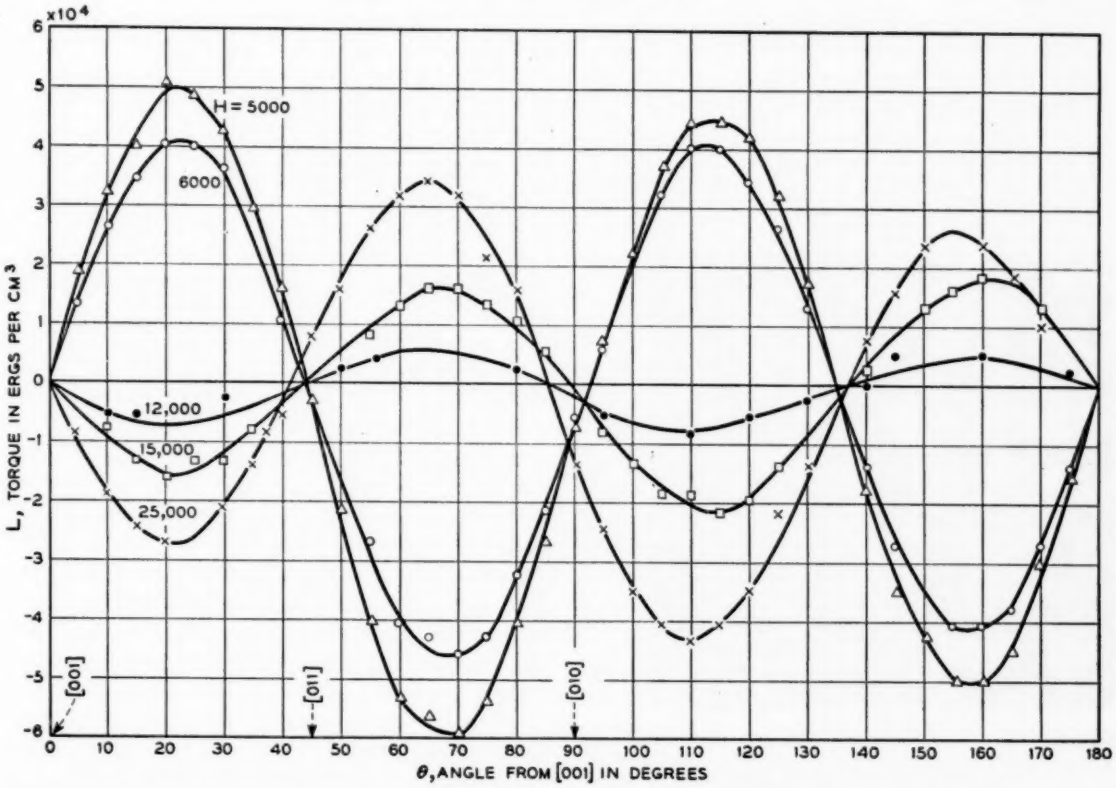


FIG. 1. Torque curves for Fe_2NiAl single-crystal disk (100), cooled from 1300°C at approximately 8°C/sec. Coercive force of disk is 614 oersteds.

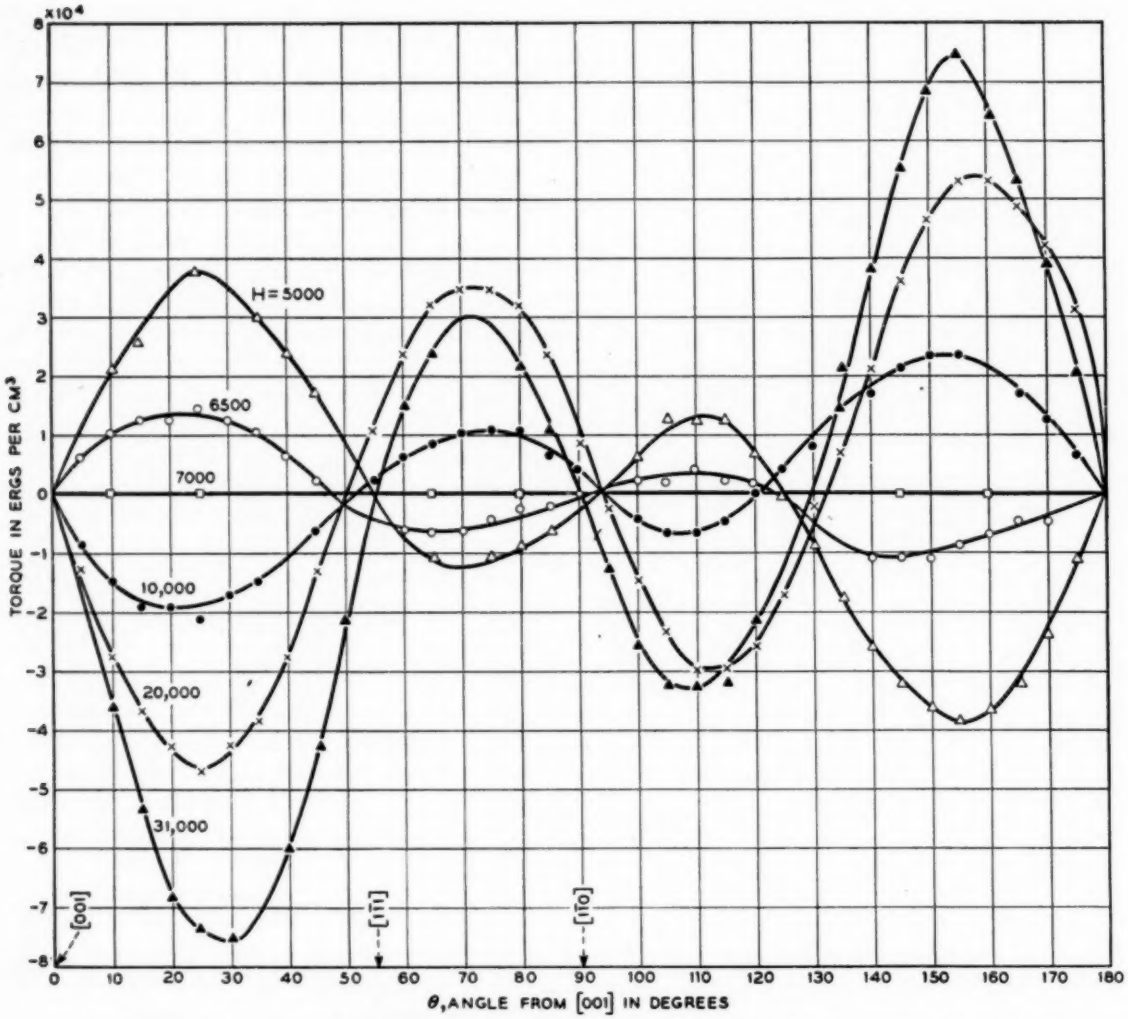


FIG. 2. Torque curves for Fe_2NiAl single-crystal disk (110), cooled from 1300°C at approximately 8°C/sec. Coercive force of disk is 620 oersteds.

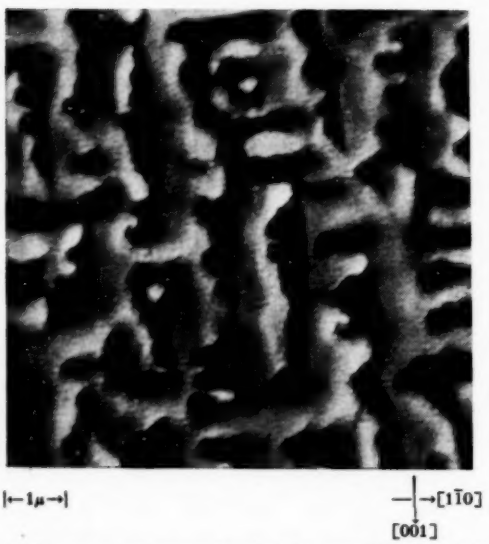


FIG. 3. Electron micrograph of single-crystal disk of Fe_2NiAl (110), cooled from 1300°C at approximately 8°C/sec, then aged at 820°C for 7 hr. Coercive force of disk is 124 oersteds.

metry as the curves for the (110) crystal disk and the amplitude of the curves decreases with increasing field.

RESULTS ON SINGLE CRYSTALS

On the basis of the results obtained on the crossed wires, the unusual behavior of the torque curves of the crystal shown in Figs. 1 and 2 is attributed to the shape anisotropy of the precipitate which is oriented in the three ⟨100⟩ directions. In the case of the crystal, however, the situation is more complicated since the torque actually reduces to zero and then changes sign. This is because there is crystal anisotropy in addition to shape anisotropy and the former has a different easy direction of magnetization from the latter. At low fields, the effect of shape predominates and at high fields crystal anisotropy is more important.

If we increase the shape effect of the precipitate by over-aging the crystal, the torque due to shape effect is predominant at all fields used. This is shown in Fig. 9 by a series of curves of apparent anisotropy constant *vs* H, obtained after subjecting the crystal to different

treatments which resulted in various coercive forces H_c. For an H_c of 600 oersteds the specimen is not over-aged and the apparent anisotropy constant which is negative decreases in magnitude to zero and finally becomes positive as H is increased. This is from the same data as are shown in Fig. 2. When the crystal is over-aged somewhat, H_c is 300 oersteds and the torque cannot be reduced to zero for values of field as high as 20 000 oersteds. Upon further overaging H_c is 47 oersteds, the shape effect is now very well developed and at an H of 20 000 the apparent anisotropy constant has the very high value of 600 000.

Figure 10 is an electron micrograph of the crystal in the latter state ($H_c=47$) and the well-developed shape anisotropy is evident. The coercive force is low in this structure because the particle size is large enough to permit domain wall motion. In contrast to this structure, Fig. 11 shows an electron micrograph of a crystal for which H_c is 525 oersteds. Here the elongation of the precipitate cannot be resolved in the electron micrograph but the behavior of the torque curves indicates that the precipitate is elongated.

The torque curves in Figs. 1 and 2 were obtained on single crystals of Fe_2NiAl in the (100) and (110) planes, respectively, when $H_c=600$ oersteds. The results show qualitatively the same effect with increasing fields. In both cases, the torque reduces with increasing field, passes through zero and then changes sign. The torque curve appears to pass through zero at somewhat higher values of H for the (100) crystal plane ($H=12\,000$) than for the (110) plane crystal ($H=7000$). This is also shown by the experimental torque curves for 2 wires (100) and 3 wires (110), the latter structure having a lower amplitude of torque at high fields than the former.

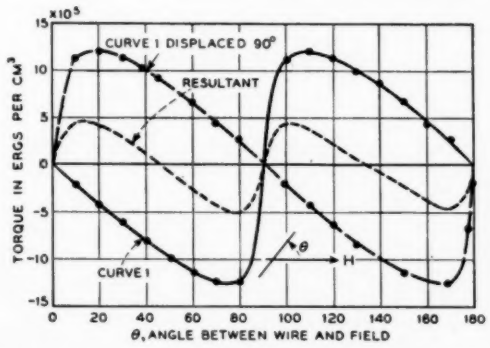


FIG. 4. Torque curve for a single Mo Permalloy wire at $H=4000$ oersteds.

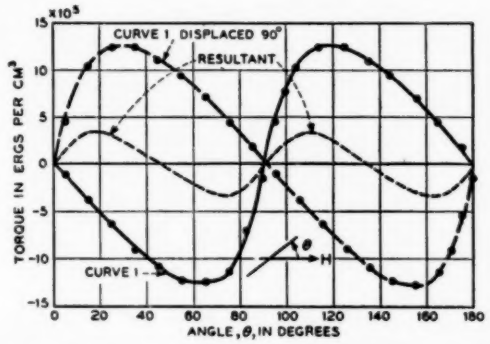


FIG. 5. Torque curve for a single Mo Permalloy wire at $H=7000$ oersteds.

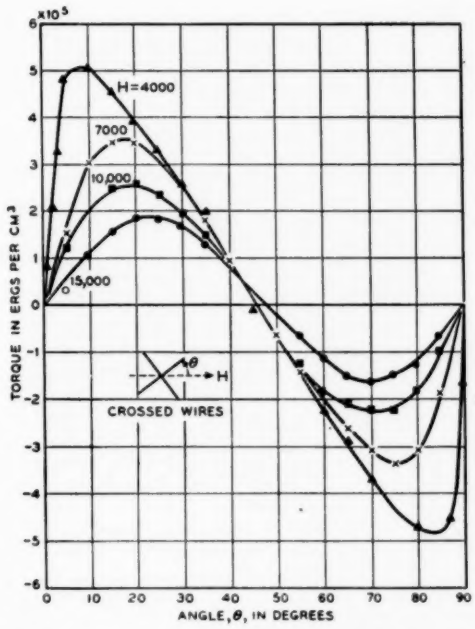


FIG. 6. Torque curves for two Mo Permalloy wires at right angles to each other.

In order to obtain the final value of the crystal anisotropy, it is necessary to extrapolate to infinite H, and when this is done values of K of 110 000 and 150 000 ergs per cm³ are obtained for the (100) and (110) planes, respectively.

The relative importance of the shape and crystal anisotropies in determining the coercive force was found by measuring the directions of easy magnetization, coercive force, and saturation magnetization at elevated temperatures. A crystal disk with a (110) face was placed in the torque magnetometer and heated to 600°C. Full torque curves were not measured at 600°C; at high field strengths ($H=18\,000$) the crystal was isotropic. The [001] direction of easy magnetization had disappeared, indicating that the crystal anisotropy had been practically reduced to zero at this temperature. At low field strengths ($H=4000$) however, substantial torques were still necessary to turn the crystal away from ⟨111⟩ directions. Therefore the coercive force of the crystal at 600°C which was found to be 320 oersteds must be due almost entirely to shape anisotropy.

To find the value of coercive force due to shape anisotropy at room temperature, a temperature run (Fig. 12) on a polycrystalline bar of Fe_2NiAl was made to determine the change in saturation with temperature. The coercive force of the bar was 320 oersteds at 600°C

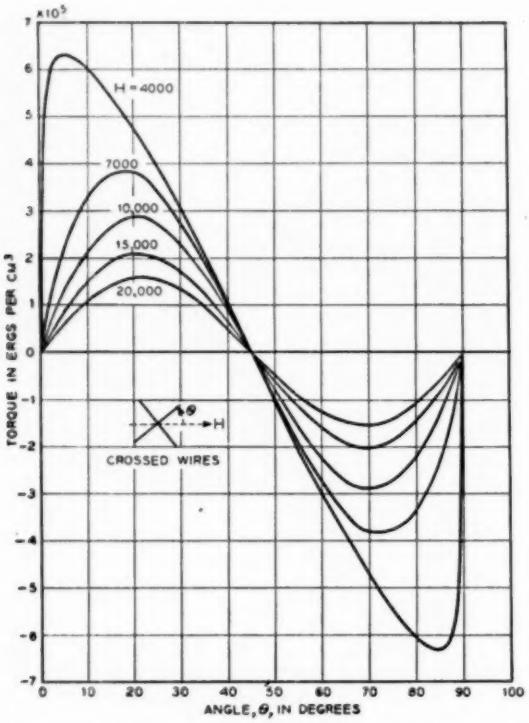


FIG. 7. Calculated torque curves for two wires at right angles to each other.

and since the coercive force due to shape anisotropy is proportional to the saturation, the room temperature value of coercive force should be 465 oersteds. This leaves approximately 155 oersteds to be accounted for by crystal anisotropy and therefore shape anisotropy accounts for about 75 percent of the coercive force.

CONCLUSIONS

As a result of this work, it appears that shape anisotropy of the fine particles is responsible for the unusual behavior of the torque curves obtained on single crystals of Fe_2NiAl. We have not been able to detect elongated particles in electron micrographs taken when the crystal is in its high coercive force state. Neither is shape anisotropy apparent, at first, from torque curves obtained on these crystals since the curves show a fourfold symmetry in the (100) plane which is indicative of cubic crystalline anisotropy. However, our analysis shows that the fourfold symmetry obtained at low values of H (about 5000) is due principally to shape anisotropy of plates (made up of rods[7]) oriented 90° to each other. The torque resulting from the shape of the plates arranged at 90° to each other eventually becomes zero at infinite H. Therefore, by extrapolation, we have been able to obtain the crystal anisotropy of the alloy. The relative importance of the crystal *versus* the shape anisotropy has been determined by making measurements at 600°C where the crystal anisotropy practically disappears. As a result of these measurements, it is concluded that shape anisotropy is a major factor in determining the coercive force of single crystals of Fe_2NiAl.

The experimental work, also, supplies us with a simple explanation for the easy direction of magnetization in these crystals. Under permanent magnet conditions and at field strengths of less than 7000 oersteds the [111] direction is easy. This is due to the shape effect of three sets of plates mutually perpendicular to each other. At higher fields, the effect of shape on the anisotropy tends

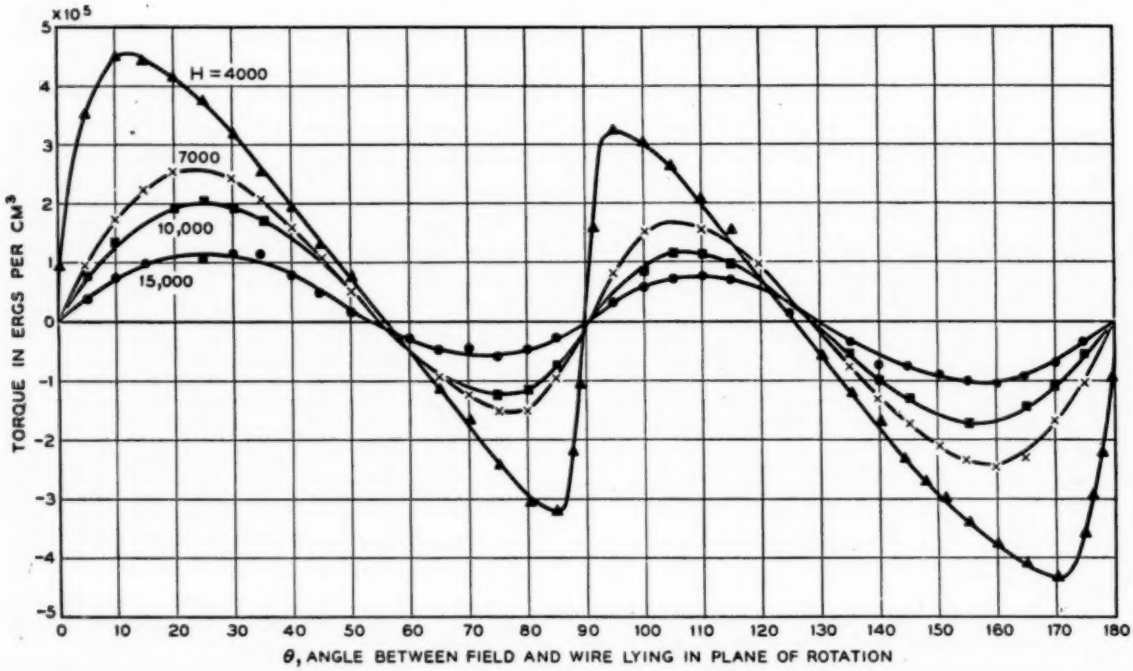


FIG. 8. Torque curves for three mutually perpendicular wires measured in the plane corresponding to (110).

to diminish and the easy direction becomes the [001] direction as a result of crystal anisotropy.

Since Fe_2NiAl is closely related to Alnico 5 and since its electron micrographs are similar, the explanations for the easy directions of magnetization in crystals of Fe_2NiAl also apply to Alnico 5, which is being investigated in a similar manner.

The anisotropy recently ascribed by Néel[4] to the surface effects of small particles can contribute to the torque in the same way as the shape anisotropy does. If surface anisotropy plays a significant role in Fe_2NiAl, we can still conclude that fine precipitated particles are present as plates. When the particles are large the surface anisotropy becomes negligible but the shape effect persists; when they are small, surface effects are present along with shape anisotropy.

ACKNOWLEDGMENTS

The authors are greatly indebted to J. K. Galt for valuable discussions during the work, to R. D. Heidenreich and Eloise Koonce for the electron micrographs and to Miss B. B. Cetlin and A. J. Williams for technical assistance. We also wish to thank D. H. Wenny for supplying the alloys for these experiments.

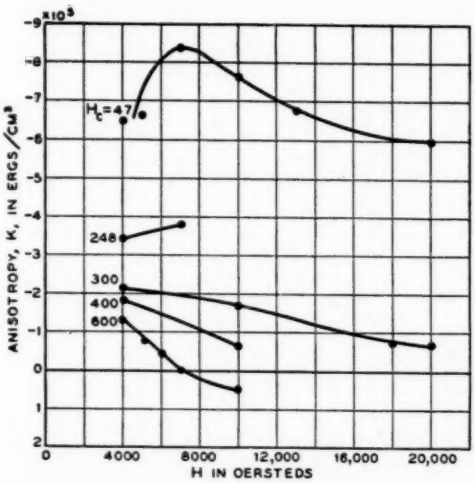


FIG. 9. Magnetic anisotropy constant K, *vs* H for (110) crystals having various coercive forces.

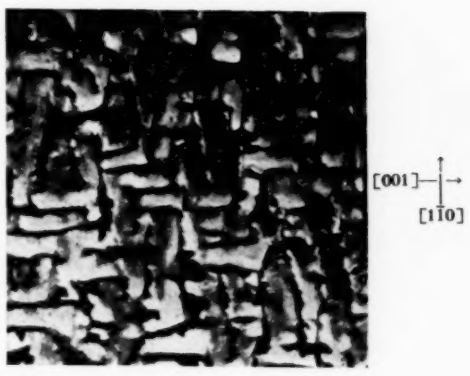


FIG. 10. Electron micrograph of single-crystal disk of Fe_2NiAl (110), cooled from 1300°C at approximately 8°C/sec, then heated at 820°C for 35 hr. Coercive force of disk is 47 oersteds.

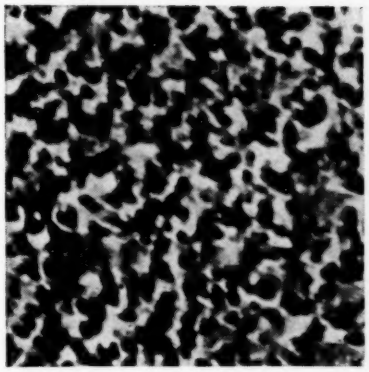


FIG. 11. Electron micrograph of single-crystal disk of Fe_2NiAl (100), cooled from 1300°C at approximately 8°C/sec. Coercive force of disk is 525 oersteds.

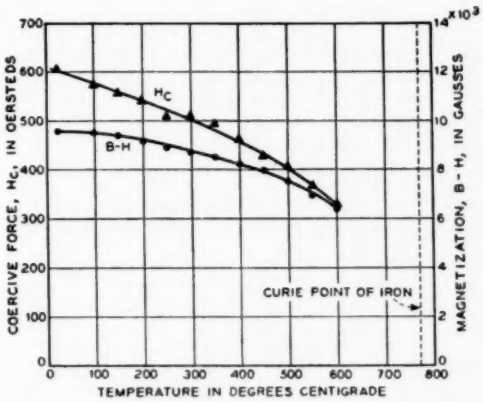


FIG. 12. Saturation magnetization and coercive force *versus* temperature on a polycrystalline bar of Fe_2NiAl.

APPENDIX I

Torque in High Fields

According to Fig. 13, the torque L per unit volume for one wire is

$$L_1 = HI_s \sin\alpha_1. \tag{1}$$

This is the torque resulting from the action of the applied field H on the magnetization of the wire; opposing it is an equal and opposite restoring torque which tends to align the domains along the length of the wire; this is equal in magnitude to

$$\begin{aligned} L_1 &= -[N_t I_s \sin(\theta-\alpha_1) - N_l I_s \sin(\theta-\alpha_1)] I_s \cos(\theta-\alpha_1) \\ &= -\tfrac{1}{2}(N_t - N_l) I_s^2 \sin 2(\theta-\alpha_1), \end{aligned} \tag{2}$$

where N_t and N_l are, respectively, the transverse and longitudinal demagnetizing factors of the wire. When the wire is completely saturated, $\alpha_1 = 0$ and the torque reduces to

$$L_1 = -\tfrac{1}{2}(N_t - N_l) I_s^2 \sin 2\theta.$$

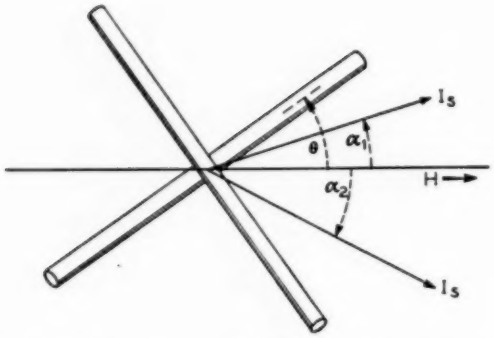


FIG. 13. Two wires crossed at right angles, simulating rods of precipitation.

The torque for the other wire at 90° to the first is opposite in sign to L_1 and is

$$L_2=\tfrac{1}{2}(N_t-N_l)I_s^2\sin2(\theta+\alpha_2).$$

When $\alpha_2=0$, the resultant torque of the 2 wires is

$$L=\tfrac{1}{2}(L_1+L_2)=0.$$

This limit is approached in very high fields.

When the wires are nearly saturated α_1 and α_2 are small and nearly equal. As a first approximation we put $\sin\alpha_2=\sin\alpha_1=\alpha_1$, and find the resultant torque per unit volume to be

$$L=\tfrac{1}{2}(L_1+L_2)=-\tfrac{1}{4}(N_t-N_l)I_s^2[\sin2(\theta-\alpha_1)-\sin2(\theta+\alpha_2)]=\alpha_1(N_t-N_l)I_s^2\cos2\theta. \quad (3)$$

From Eqs. (1) and (2), at all values of α_1

$$HI_s\sin\alpha_1=\tfrac{1}{2}(N_t-N_l)I_s^2\sin2(\theta-\alpha_1). \quad (4)$$

When α_1 is small,

$$\alpha_1=\tfrac{1}{2}(N_t-N_l)I_s(\sin2\theta)/H \quad (5)$$

for the one wire, and a similar expression in α_2 applies

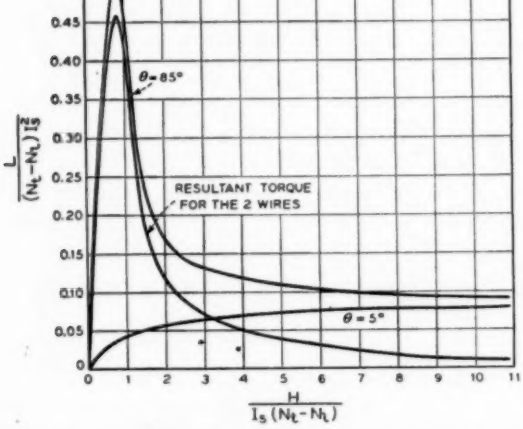


FIG. 14. Curves showing how the torque varies with H for two complementary values of θ. The resultant torque for two perpendicular wires is ½ of that shown by the difference curve.

to the other wire. The torque per unit volume for the pair of wires is then, from (3) and (5)

$$L=\tfrac{1}{4}(N_t-N_l)^2I_s^3(\sin4\theta)/H, \quad (6)$$

which approaches zero as H becomes indefinitely large. This relation is in accord with the period and amplitude of the experimentally determined curves on crossed wires (Fig. 6).

APPENDIX II

Torque in Field of Any Strength

From Eqs. (1) and (2) of Appendix I,

$$\frac{L_1}{(N_t-N_l)I_s^2}=\frac{\sin2(\theta-\alpha_1)}{2},$$

and

$$\frac{H}{(N_t-N_l)I_s}=\frac{\sin2(\theta-\alpha_1)}{2\sin\alpha_1}.$$

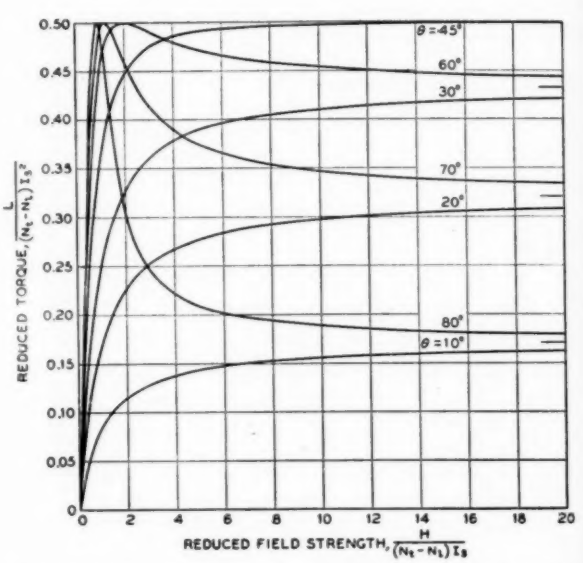


FIG. 15. Curves showing torque vs field-strength for particles having various transverse and longitudinal demagnetizing factors N_t and N_l, and saturation I_s.

A value of θ is selected, and these two expressions are calculated for this θ and various values of α_1 and a curve plotted as shown in Fig. 14. More specifically $L_1/(N_t-N_l)I_s^2$ is plotted as a function of $H/(N_t-N_l)I_s$ with θ constant. The resultant torque, $L/(N_t-N_l)I_s^2$, of two crossed wires is half the difference in height of the ordinates of the curves for θ and $90°-\theta$. From a series of such curves, curves of the type of Fig. 14 can be obtained directly. The required value of H is selected, I is determined by the material, and N_t and N_l by the physical dimensions of the wire. In the experiments performed $I_s=690$, $N_t=6.226$, and $N_l=0.113$.

Curves for various other values of θ are given in Fig. 15. Similar calculations and experiments have been carried out for three mutually perpendicular wires.

An Electrodynamic Perturbation Theorem, with Application to Nonreciprocal Systems

ALFRED G. REDFIELD*
Gordon McKay Laboratory, Harvard University, Cambridge, Massachusetts
(Received November 16, 1953)

By means of a slight modification of the macroscopic electrodynamic reciprocity theorem it is possible to obtain the perturbation of the admittance matrix of an electrodynamic device due to a suitable perturbation of the dielectric susceptibility, conductivity, and magnetic susceptibility tensors in the device. The theory is applicable to nonreciprocal systems, but in practice can be used to obtain only the first-order perturbation of the admittance matrix. An expression relating the current induced in part of a system due to the motion of current in another part of the system is similarly obtained. Dc and microwave gyrators are treated, and the connection is pointed out between Faraday rotation in a transmission cavity and the magnetic energy splitting of its circularly polarized normal modes.

INTRODUCTION

THE macroscopic electrodynamic reciprocity theorem[1,2] was first proven by Carson[3] and by Pleijel[3] for systems in which radiation is important. As Carson pointed out,[4] his treatment is not valid for the case of transmission through an ionized gas in a dc magnetic field. Recently, several authors[5,6] have pointed out other mechanisms leading to nonreciprocity in electrodynamic systems, and several useful nonreciprocal devices have been build.[7,8] The nonreciprocity of such a device is a consequence of the asymmetry of the dielectric constant, conductivity, and magnetic susceptibility tensors of matter in the device, brought about by application of a large magnetic field. The proof of the reciprocity theorem requires that these tensors be symmetric.

In this paper a useful modification of the reciprocity theorem is derived for a generalized system. This theorem is used to evaluate the antisymmetric part of the admittance matrix of an electrodynamic system in terms of the electromagnetic fields produced by the applied voltages in the system, and to obtain the perturbation of the admittance matrix produced by suitable perturbations of the system. These results are applicable as long as the macroscopic Maxwell equations are, and can be used to treat dc and microwave gyrators, as well as reciprocal systems. The main limitation to the method is that it can only be carried out to a first order of approximation. It seems likely that similar methods can be applied to mechanical, electromechanical, and hydrodynamic systems.

GENERAL THEOREMS

Consider a system C' which is bounded but not necessarily closed. We consider steady-state solutions of frequency ω, and all field quantities can be complex according to the usual ac convention. The system is assumed to obey the macroscopic Maxwell equations (MKS units):

$$\nabla\times\mathbf{H}'-i\omega\mathbf{E}'-\mathbf{u}'=0, \tag{1'}$$

$$\nabla\times\mathbf{E}'+i\omega\mu_0(\mathbf{H}'+\mathbf{M}')=0, \tag{2'}$$

where $\mathbf{u}$ is the total real current (conduction plus dielectric polarization but not displacement), $\mathbf{M}$ is the magnetization, and the other symbols are conventional. The current and magnetization are given by

$$\mathbf{u}'=(i\omega\chi_e\epsilon_0+\sigma)(\mathbf{E}'+\mathbf{F}')\equiv\lambda(\mathbf{E}'+\mathbf{F}'), \tag{3'}$$

$$\mathbf{M}'=\chi_m(\mathbf{H}'+\mathfrak{F}'), \tag{4'}$$

where χ_e, σ, and χ_m are the electric susceptibility, conductivity, and magnetic susceptibility tensors (which can be nonsymmetric), and $\mathbf{F}'$ and $\mathfrak{F}'$ are the electromotive and magnetomotive forces or sources which give rise to the electric and magnetic fields in the system. (This definition of emf and mmf is rather arbitrary but is easily transformed to any other definition.) The tensor λ is introduced for convenience.

We also consider a second system C'' in which the fields also obey Maxwell's equations and are set up by a second distribution of forces $\mathbf{F}''$ and $\mathfrak{F}''$, characterized

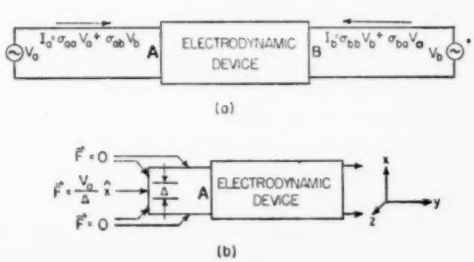


FIG. 1. Four pole device.

* National Science Foundation Postdoctoral Fellow.
[1] S. A. Schelkunoff, *Electromagnetic Waves* (D. Van Nostrand Company, Inc., New York, 1943), p. 476.
[2] Montgomery, Dicke, and Purcell, *Principles of Microwave Circuits* (McGraw-Hill Book Company, Inc., New York, 1948), p. 141.
[3] J. R. Carson, Bell System Tech. J. **3**, 393 (1924); **9**, 325 (1930).
[4] J. R. Carson, Proc. Inst. Radio Engrs. **17**, 952 (1929).
[5] H. Suhl and L. R. Walker, Phys. Rev. **86**, 122 (1952).
[6] E. M. McMillan, J. Acoust. Soc. Am. **18**, 344 (1946).
[7] C. L. Hogan, Bell System Tech. J. **31**, 1 (1952).
[8] Mason, Hewitt, and Wick, J. Appl. Phys. **24**, 166 (1953).

by the equations

$$\mathbf{u}''=(\lambda+\delta\lambda)(\mathbf{E}''+\mathbf{F}''), \tag{3''}$$

$$\mathbf{M}''=(\chi_m+\delta\chi_m)(\mathbf{H}''+\boldsymbol{\beta}''), \tag{4''}$$

where $\delta\lambda$ and $\delta\chi_m$ are defined analogously to λ and χ_m, and express the difference between systems C' and C''. The field quantities in Eq. (3'') and (4'') are regarded as functions of the same space coordinate $\mathbf{x}$ as those in Eqs. (3') and (4'). Then the following identity holds:

$$\int_s(\mathbf{E}'\times\mathbf{H}''-\mathbf{E}''\times\mathbf{H}')\cdot d\mathbf{s}$$

$$+\int_v[\mathbf{u}'\cdot\mathbf{F}''-\mathbf{u}''\cdot\mathbf{F}'-i\omega\mu_0(\mathbf{M}'\cdot\boldsymbol{\beta}''-\mathbf{M}''\cdot\boldsymbol{\beta}')]dv$$

$$=\int_v[i\omega\mu_0(\mathbf{H}'+\boldsymbol{\beta}')(2\chi_m{}^a+\delta\chi_m)(\mathbf{H}''+\boldsymbol{\beta}'')$$

$$-(\mathbf{E}'+\mathbf{F}')(2\lambda^a+\delta\lambda)(\mathbf{E}''+\mathbf{F}'')]dv, \tag{5}$$

where $\chi_m{}^a$ and λ^a are the antisymmetric parts of χ_m and λ, and the surface integration is over the surface of the region of the volume integration, which is arbitrary. The derivation of this equation is simple and very similar to that of reciprocity theorem[1,2] and is therefore omitted.

To apply Eq. (5) to get the asymmetry and perturbation of the admittance matrix σ_{ij} of an electrodynamic device (Fig. 1a) the integration of Eq. (5) is taken over all space and the terminal loops A and B are assumed equivalent to plain loops in which $|\mathbf{F}|=V_a/\Delta$ and V_b/Δ for an infinitesimal length Δ, and in which $\mathbf{F}$ points along the wire at the point where Δ is located (Fig. 1b). The surface integral vanishes at infinity[2] and one easily obtains

$$\sigma_{ij}'-\sigma_{ji}'=\int[i\omega\mu_0\mathbf{H}_j'(2\chi_m{}^a+\delta\chi_m)\mathbf{H}_i''$$

$$-\mathbf{E}_j'(2\lambda^a+\delta\lambda)\mathbf{E}_i'']dv, \tag{6}$$

where $\mathbf{E}_j'$, $\mathbf{H}_j'$ are the fields in the unperturbed device (characterized by λ and χ_m) when $V_j=1$, $V_k=0$ $(k\neq j)$, and $\mathbf{E}_i''$, $\mathbf{H}_i''$ are the fields in the perturbed device (characterized by

$$\lambda+\delta\lambda,\ \chi_m+\delta\chi_m) \quad \text{when} \quad V_i=1,\ V_k=0\ (k\neq i).$$

Equation (6) can easily be generalized to include magnetic and waveguide inputs. In the latter case the region of integration is taken to exclude the interior of the microwave generator and the result is identical to Eq. (6), under the waveguide voltage and current conventions of reference 2. If the unperturbed system is reciprocal, Eq. (6) can be used to obtain the change $\delta\sigma_{ij}$ due to the perturbation. If the perturbation is zero, Eq. (6) gives the antisymmetric part of the admittance matrix. The preceding results are obviously applicable to devices having more than two terminals.

The current induced in an external current loop (or magnetic loop or wave guide) by a given distributed emf or current (or mmf or magnetization) in a reciprocal device can be similarly obtained:

$$I_i=\int(\mathbf{F}\cdot\lambda\mathbf{E}_i-i\omega\mu_0\boldsymbol{\beta}\cdot\chi_m\mathbf{H}_i)dv$$

or

$$I_i=\int(\mathbf{u}_g\cdot\mathbf{E}_i-i\omega\mu_0\mathbf{M}_g\cdot\mathbf{H}_i)dv, \tag{7}$$

where $\mathbf{E}_i$ and $\mathbf{H}_i$ are the fields in the system where $V_i=1$, $\boldsymbol{\beta}=0$ and $\mathbf{F}=0$, and $\mathbf{u}_g$ and $\mathbf{M}_g$ are the impressed current and magnetization (which are equivalent to the impressed emf $\mathbf{F}$ and mmf $\boldsymbol{\beta}$ and related to them by λ and χ_m) which produce the current I_i in loop i. Equation (7) is a generalization of a theorem proven by Schockley.[9]

APPLICATIONS

All the equations derived so far have been exact but have required knowledge of the fields in both the perturbed and unperturbed systems. In applying these formulas it is necessary to approximate these fields by the fields in the unperturbed system. There is no obvious way to introduce a higher order correction into the theory, and this is a serious limitation in microwave problems involving ferrites, where the field perturbations are likely to be large. Many problems to which the theory is applicable can also be treated in other ways. Nevertheless, this theory may be useful for some future application and can perhaps give more physical insight into the effects of various perturbations than can other theoretical treatments.

At low frequencies $\chi_m{}^a$ is finite or zero and $\lambda^a\mathbf{E}=\mu_H\mathbf{E}\times\mathbf{H}_0$ where μ_H is the Hall mobility and H_0 is the dc magnetic field. Therefore the Hall voltage δv_b which would appear on terminals B if they were open-circuited, due to an applied voltage V_a, for an arbitrary arrangement of electrodes on a conductor, is given by

$$\frac{\delta v_b}{V_a}=\frac{\delta\sigma_{ba}}{\sigma_{bb}}\simeq\frac{\int\mu_H\sigma(\mathbf{E}_b\times H_0)\cdot\mathbf{E}_a dv}{\int\mathbf{E}_b\cdot\sigma\mathbf{E}_b dv}. \tag{8}$$

This equation can be used to analyze the author's method[10] of measuring the electronic Hall mobility in insulating photoconductors. It apparently holds even in the presence of space charge effects, provided they are linear. The dc case has also been treated by Wick.[11]

[9] W. Shockley, J. Appl. Phys. **9**, 635 (1938).
[10] A. G. Redfield, Phys. Rev. **91**, 373 (1953).
[11] Reference 8, footnote 4.

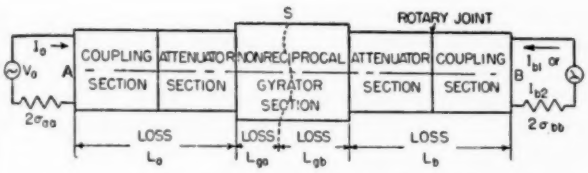


FIG. 2. Generalized microwave gyrator. The dotted line S denotes a surface which cuts completely across the gyrator.

For microwave gyrator applications it is useful to consider the arrangement of Fig. 2. The gyrator section may be a waveguide or cavity containing a ferrite or an ionized gas in a dc magnetic field, and the attenuator and coupling sections are assumed reciprocal. The gyrator and attenuator sections are assumed to be invariant under a 90° rotation about the horizontal axis so that there are two similar orthogonal modes of transmission (which we will call 1 and 2) in each direction. Mode 1 is defined as the mode excited by loop A when the dc magnetic field and magnetization are zero, and the section coupling loop B can be rotated 90° about the horizontal axis with respect to the rest of the system so that it is coupled with either mode 1 or mode 2. I_{b1} denotes the current in loop B when the coupling section is in position 1; I_{b2} denotes the current in loop B when the coupling section is in position 2. I_{b2} is zero unless the dc magnetic field induces coupling between the two modes, and its magnitude is proportional to the resulting rotation of the microwaves in the gyrator. The coupling sections present admittances $2\sigma_{aa}$ and $2\sigma_{bb}$ to the loops A and B, so that the generators see admittances σ_{aa} and σ_{bb}. Without loss of generality we can assume that the coupling and attenuator sections have been designed so that σ_{aa} and σ_{bb} are real, and that I_{b1} has the same phase as V_a when the magnetic field is off. L_a and L_b are assumed large so that σ_{aa} and σ_{bb} do not depend significantly on the rotation of the coupling section B with respect to the system.

We will regard the gyrator as unperturbed when the dc magnetic field is off (λ^a, $\chi_m{}^a=0$); application of the magnetic field introduces a nonreciprocal perturbation into the system.

Using Eq. (6) we have

$$I_{b2}=V_a\sigma_{ba2}$$

$$=V_a\int(i\omega\mu_0\mathbf{H}_a'\chi_m{}^a\mathbf{H}_{b2}''-\mathbf{E}_a'\lambda^a\mathbf{E}_{b2}'')dv, \quad (9)$$

where σ_{ba2}, $\mathbf{E}_{b2}''$, and $\mathbf{H}_{b2}''$ are σ_{ba}, $\mathbf{E}_b''$ and $\mathbf{H}_b''$ defined as before and evaluated with the loop B in position 2. For small gyrations $\mathbf{E}_a'$, $\mathbf{H}_a'$, $\mathbf{E}_b''$, and $\mathbf{H}_b''$ can be approximated by their values in the unperturbed system.

Now assume that the coupling section is in position 1. Generator V_a supplies power $\frac{1}{2}V_a{}^2\sigma_{aa}$ to the system. Half of this power is dissipated in the external matching admittance $2\sigma_{aa}$. Only a fraction $L_aL_{ga}L_{gb}L_b$ of the remaining power is dissipated in the external admittance $2\sigma_{bb}$:

$$\tfrac{1}{2}I_{b1}{}^2/2\sigma_{bb}=\tfrac{1}{4}V_a{}^2\sigma_{aa}L_aL_{ga}L_{gb}L_b, \quad (10)$$

so

$$I_{b2}/I_{b1}=\sigma_{ba2}/(\sigma_{aa}\sigma_{bb}L_aL_bL_{ga}L_{gb})^{\frac{1}{2}}. \quad (11)$$

The power crossing the surface S is given by $\frac{1}{4}V_a{}^2\sigma_{aa}L_aL_{ga}$. Equating this expression with that given by integrating the Poynting vector normal component over S gives

$$\sigma_{aa}L_aL_{ga}=2\,\mathrm{Re}\int_s(\mathbf{E}_a\times\mathbf{H}_a{}^*)\cdot d\mathbf{s}. \quad (12)$$

Applying similar reasoning to eliminate $\sigma_{bb}L_bL_{gb}$, gives

$$\frac{I_{b2}}{I_{b1}}\equiv\Theta\simeq$$

$$\frac{\int(i\omega\mu_0\mathbf{H}_a\cdot\chi_m{}^a\mathbf{H}_{b2}-\mathbf{E}_a\cdot\lambda^a\mathbf{E}_b)dv}{2\left[\mathrm{Re}\int_s(\mathbf{E}_a\times\mathbf{H}_a{}^*)\cdot d\mathbf{s}\;\mathrm{Re}\int_s(\mathbf{E}_{b2}\times\mathbf{H}_{b2}{}^*)\cdot d\mathbf{s}\right]^{\frac{1}{2}}}. \quad (13)$$

In cases where Θ is pure real this expression gives the angle of rotation of the gyrator (provided the modes of transmission in the gyrator are only onefold symmetric about the axis, i.e., like a TE_{11} mode); if Θ is complex the output of the gyrator is elliptically polarized. If the gyrator is cylindrically symmetric this formula can presumably be extended to large rotations by iteration, as long as the rotation per wavelength is much smaller than one radian. Equation (13) does not depend on the details of the coupling; all that is required is to know the fields in the gyrator for one mode of propagation. Frequently $\mathbf{E}_{b2}$ and $\mathbf{H}_{b2}$ can be obtained from $\mathbf{E}_a$ and $\mathbf{H}_a$ by symmetry. Treatment of higher modes in cylindrical gyrators can be accomplished by methods similar to the above. Waveguide gyrators have also been treated by Suhl and Walker.[5]

If the gyrator is a transmission cavity near resonance, it is simpler to use the relation $\Theta=\sigma_{ba2}/\sigma_{ba1}$. From conservation of energy and because $\sigma_{ab1}\simeq\sigma_{ba1}$, and $\sigma_{ba1}\simeq\sigma_{ba1}{}^*$, it is easy to show that $\sigma_{ba1}\simeq\mathrm{Re}\int(\mathbf{E}_a{}^*\cdot\lambda\mathbf{E}_{b1}-i\omega\mu_0\mathbf{H}_a{}^*\cdot\chi_m\mathbf{H}_{b1})dv$, where the integrand (near resonance) is negligible except inside the cavity. Assuming the cavity is axially symmetric and is symmetric about a plane perpendicular to its axis, it follows that

$$\Theta=(Q/\omega\epsilon)\int(i\omega\mu_0\mathbf{H}_1\chi_m{}^a\mathbf{H}_2-\mathbf{E}_1\lambda^a\mathbf{E}_2)dv, \quad (14)$$

where Q is the Q of the cavity, ϵ is the energy stored in the cavity for fields $\mathbf{E}_1$, $\mathbf{H}_1$ in the cavity, $\mathbf{E}_1$ and $\mathbf{H}_1$ are the unperturbed fields in the cavity at resonance, and $\mathbf{E}_2$ and $\mathbf{H}_2$ are obtained from $\mathbf{E}_1$ and $\mathbf{H}_1$ by a 90° rotation. To get Eq. (14), we have used the fact that σ_{ba1} is, by the symmetry of the fields, approximately

equal to the average power dissipated in the cavity by one mode alone. Equation (14) is valid as long as the magnetic splitting of the resonance frequency in the cavity is small compared to the resonance width. This is equivalent to the assumption that $\Theta \ll 1$, as will be shown below. The effect of generalized media in cavities has also been treated by Berk and Lax.[12]

Faraday rotation in a transmission cavity can be related to the splitting of the degeneracy and the change in Q of the cavity in a simple and instructive way. We assume that at time $t=0$, we somehow excited a linearly polarized mode of oscillation in a cavity; this linearly polarized mode can be regarded as the sum of two oppositely rotating circularly polarized modes.[13] If the degeneracy of the two modes is split by a magnetic field, the resulting oscillation will look like a plane oscillation whose plane of polarization is slowly rotating with an angular frequency equal to one half the angular frequency difference between the two circular modes $\Delta\omega$. Such an oscillation will be damped out in a mean time $\tau=2Q/\omega$, so that the angle through which such a vibration will turn before being damped out is $Q\Delta\omega/\omega$. Since the input waveguide is continually exciting these two circular modes in equal amplitudes (assuming $\Delta\omega \ll \omega/Q$) and the output waveguide is sensitive to the mean polarization of the resulting precessing and decaying oscillation in the cavity, the output waveguide polarization will be oriented at an angle $\Theta=Q\Delta\omega/\omega$ relative to the input.

To make this derivation slightly more general we assume that the two circular modes have a time dependence

$$A^+=\exp(i\omega_1 t-\omega_1 t/Q_1) \text{ and } A^-=\exp(i\omega_2 t-\omega_2 t/Q_2).$$

If these modes are excited equally by an incoming linear mode of the waveguide at frequency $\omega=\frac{1}{2}(\omega_1+\omega_2)$ polarized in direction 1, the oscillation in the 1 direction in the cavity will have a mean amplitude and phase relative to the incoming wave given by

$$C\int_0^\infty (A^++A^-)\exp(-i\omega t)dt,$$

where C is a coupling constant. The mean amplitude in the 2 direction will be

$$C\int_0^\infty i(A^+-A^-)\exp(-i\omega t)dt.$$

The complex angle of rotation Θ is equal to the first integral divided by the second:

$$\Theta \simeq [2Q(\omega_1-\omega_2)/(\omega_1+\omega_2)] - i[(Q_1-Q_2)/(Q_1+Q_2)], \quad (15)$$

provided $\Theta \ll 1$. This equation indicates that a Faraday rotation in a cavity should be detectable even when a resonance splitting is not.

[12] A. D. Berk and B. Lax, Quart. Prog. Rep., Mass. Inst. Tech. Res. Lab. of Electronics, 1952 and 1953.

[13] J. O. Artman and P. E. Tannenwald, Phys. Rev. **91**, 1014 (1953).

JOURNAL OF APPLIED PHYSICS VOLUME 25, NUMBER 8 AUGUST, 1954

Short-Time Frequency Measurement of Narrow-Band Random Signals in the Presence of Wide-Band Noise*

PETER M. SCHULTHEISS, CONRAD A. WOGRIN, AND FELIX ZWEIG
Department of Electrical Engineering, Yale University, New Haven, Connecticut

(Received October 1, 1953)

In certain applications it is necessary to measure the center frequency of a symmetrical power spectrum in the shortest possible time. Two instrumentations for short-time frequency measurement, the autocorrelator and the frequency discriminator, are examined for their ability to accomplish this task. The analysis compares the performance of the two systems and indicates the effect of such factors as signal-to-noise ratio and filter adjustment.

INTRODUCTION

IN the measurement of physical quantities by electrical means, it is not uncommon to translate the actual information into frequency shift of a carrier wave. The process is essentially one of frequency modulation; the information can be recovered at any time by virtue of the known one to one relation between the signal's "instantaneous frequency" and the desired physical data. While the term "instantaneous frequency" is self-contradictory in a sense, this point offers no serious obstacle to formal mathematical analysis. The problem of frequency modulation in fact has been treated extensively and rigorously in the literature.[1]

The simple process described thus far is often complicated by the very manner in which the data are taken. If the original information is, for instance, supplied from a source subject to random disturbances, one cannot expect the continuous exact correspondence between desired information and frequency postulated above. If the measured quantity were to remain fixed, one would not receive a single frequency in that case but rather a power spectrum presumably symmetrical about some mean frequency which could be calibrated to furnish the required data. The width of this power spectrum would vary inversely with the quality of the basic measuring process. A perfect process would result in a spectrum of zero width, a single frequency, but systems of considerable practical importance yield spectra with a band width equal to 10 percent or 20 percent of their mean frequency.

The power spectrum differs in one important respect from single frequency data. The frequency of a sine wave can be determined precisely in a relatively short time interval, for instance, by measuring the time between two successive zeros and assuming periodicity. To establish the mean frequency of a power spectrum, on the other hand, an infinite length of time is required. Observation over any finite period can only determine the true (long-time) mean frequency to within certain statistical bounds. Since the physical quantity under observation will generally vary as a function of time, very definite limits are imposed on the permissible observation period and the accuracy attainable in a given length of time thus becomes exceedingly important. It is the purpose of this discussion to compare various basic techniques of frequency measurement with a view towards maximizing the accuracy attainable in a predetermined time interval.

The frequency sensitive element may take a variety of forms. It may consist of a conventional frequency discriminator, essentially a device which measures two points on the power spectrum and deduces the mean frequency from symmetry considerations. Alternatively one may use an autocorrelator whose operation might be described loosely as the time domain analog of the power spectrum measurement performed by the discriminator. Still another method employs a unit recording the number of zeros or maxima occurring in a given length of time. The present paper is concerned only with the first two instrumentations because they lend themselves to closely similar analytical treatments while the third is approached more conveniently from a somewhat different point of view.

A FIGURE OF MERIT FOR SHORT TIME FREQUENCY MEASUREMENT

The operation of a frequency sensitive device must clearly depend on the statistical properties of the voltage applied to its input terminals. For the purposes of this paper it will be assumed that the input possesses a normal amplitude distribution, so that it might be thought of as thermal noise after passage through a narrow-band filter. This is the simplest possible assumption, and one which probably comes close to the truth for an important class of systems. The actual shape of the power spectrum is not very critical as long as the center frequency alone is of interest. It is only necessary that it be narrow compared to the center frequency and symmetrical about it. A Gaussian characteristic (Fig. 1) meets both of these requirements and has the additional advantage of leading to relatively simple mathematical manipulations.

Before various techniques for short-time frequency measurement can be compared, it becomes necessary to

* The contents of this paper are taken in essence from a dissertation by Peter M. Schultheiss presented to the faculty of the Graduate School of Yale University in partial fulfilment of the requirements for the Degree of Doctor of Philosophy.

[1] See, for instance, S. Goldman, *Frequency Analysis, Modulation and Noise* (McGraw-Hill Book Company, Inc., New York 1948).

FIG. 1. The input power spectrum.

establish a figure of merit for their performance. If the physical quantity being monitored, represented by the center frequency f_0 of the power spectrum, is expected to fluctuate over a wide range, it may well be impracticable to devise a frequency sensitive element with sufficient linear range, and one will have to resort to a feedback arrangement such as shown in Fig. 2. This procedure becomes almost imperative if the narrow-band signal is superimposed on a background of wide-band noise, for any noise-limiting filter must clearly be centered about f_0, which implies either an adjustable filter or an essentially fixed f_0. Either approach requires a feedback scheme but in practice it would probably be far more cumbersome to arrange for a frequency variant filter than to build the relatively simple circuit of Fig. 2 which keeps the mean frequency at the input to the frequency sensitive element essentially fixed during normal operation.

In Fig. 2 the frequency sensitive element operates effectively as a null seeking device: If the dc loop gain[2] is large because of high amplifier gain at dc, the steady-state value of e_0 must at all times be very close to zero while the steady-state value of the output is simply related to the center frequency f_0. It appears reasonable to define as the "best" system for short time frequency measurement that which provides the maximum ratio of useful output (indication of f_0) to unwanted disturbance (from noise and the random nature of the signal). In line with this requirement it is apparent that the frequency sensitive device should furnish as large as possible a change in steady-state output per unit change in f_0 while minimizing the rms value of the output fluctuation. Designating as μ_{11_0} the variance of e_0 when $\bar{e}_0=0$, and as S the sensitivity to frequency changes for the same operating point (i.e., $S=\partial\bar{e}_0/\partial f_0|_{\bar{e}_0=0}$) one arrives at the figure of merit $(\mu_{11_0}/S)^{\frac{1}{2}}$ which diminishes as the quality of the instrumentation improves. For mathematical reasons it will actually be more convenient to work with the square of this ratio, the quantity μ_{11_0}/S^2, which will be used throughout the remainder of the paper.[3]

In summary, the present investigation is concerned with the relative merits of two devices [(a) the autocorrelator, (b) the frequency discriminator] for short-time measurement of the mean frequency of a power spectrum. Both will be treated as isolated units; the feedback loop of Fig. 2 was introduced only to show that the condition $\bar{e}_0=0$ and small excursions of f_0 can be considered as typical. Under these conditions the ratio μ_{11_0}/S^2 furnishes an intuitively appealing and mathematically justifiable figure of merit for comparison of the two instrumentations. Extension of the argument to values of $\bar{e}_0$ other than zero offers little theoretical difficulty but further complicates already cumbersome algebraic manipulations.

The next step must now be the derivation of general expressions for the output variance and sensitivity of the two frequency sensitive devices.

THE SHORT-TIME AUTOCORRELATOR

Autocorrelation and spectral density in the normal sense (averaged over an infinite interval) are entirely equivalent concepts: One can be obtained from the other by the well-known Wiener-Khintchine relation. The short-term equivalents of the two quantities on the other hand are not necessarily identical in information content.[4] It appears pertinent, therefore, to investigate both the short-time autocorrelator and the standard frequency discriminator and to compare their performance under the specified conditions.

The term short-time autocorrelation $[\psi_t(\tau)]$ will be used to designate the output of the device shown in Fig. 3. Its average output is, of course, simply the true autocorrelation $\psi(\tau)$ of the input signal so that the sensitivity computation offers no difficulty. The output variance is most readily obtained by integrating the continuous portion of the output power spectrum $W_0(f)$.

$$\mu_{11}=\int_0^\infty W_0(f)df. \tag{1}$$

The problem therefore reduces to the determination of $W_0(f)$, a task readily accomplished with a method developed by S. O. Rice.[5]

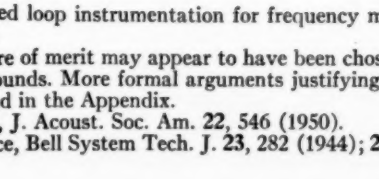


FIG. 2. Closed loop instrumentation for frequency measurement.

[2] "Dc gain" should be interpreted as the loop gain measured by breaking the loop at a point such as e_0, inserting a dc voltage and measuring the return around the loop to the point of the break.

[3] The figure of merit may appear to have been chosen on rather intuitive grounds. More formal arguments justifying its selection are contained in the Appendix.

[4] U. Fano, J. Acoust. Soc. Am. **22**, 546 (1950).

[5] S. O. Rice, Bell System Tech. J. **23**, 282 (1944); **24**, 46 (1945).

If the input signal $V(t)$ is normally distributed, it may be represented by means of a Fourier series

$$V(t)=\sum_{m=1}^{M\to\infty} C_m \cos(\omega_m t-\phi_m), \tag{2}$$

where $C_m^2\equiv 2w(f_m)\Delta f$, $\omega_m\equiv 2\pi f_m$, $f_m\equiv m\Delta f$, and $w(f)$ is the power spectrum of $V(t)$. Δf is a small increment of frequency and $\phi_1 \cdots \phi_m$ are independent random phase angles.

Then

$$\begin{aligned} V(t)V(t-\tau)&=\sum_{m=1}^{\infty}\sum_{n=1}^{\infty} C_m C_n \cos(\omega_m t-\phi_m)\\ &\quad\times\cos(\omega_n t-\omega_n\tau-\phi_n)\\ &=\frac{1}{2}\sum_{m=1}^{\infty}\sum_{n=1}^{\infty} C_m C_n\{\cos[(\omega_m-\omega_n)t\\ &\quad-\phi_m+\phi_n+\omega_n\tau]\\ &\quad+\cos[(\omega_m+\omega_n)t-\phi_m-\phi_n-\omega_n\tau]\}. \end{aligned} \tag{3}$$

The first term of Eq. (3) contains modulation products of the difference type, the second term modulation products of the sum type. With a narrow-band input centered at a relatively large frequency (f_0) only the former can contribute significant power near zero frequency. The sum terms will contribute a spectrum centered about $2f_0$ which will be effectively eliminated by the low-pass filter. However, in the interest of greater generality both terms will be retained as long as the ensuing mathematical complexity does not become prohibitive.

The difference terms contribute power in the frequency range f_k, $f_k+\Delta f$ when $m=k+l$, $n=l$ or $n=k+l$, $m=l$. For a particular value of l the two terms in question are

$$\begin{aligned} &\tfrac{1}{2}C_{k+l}C_l\{\cos[(\omega_{k+l}-\omega_l)t-\phi_{k+l}+\phi_l+\omega_l\tau]\\ &\qquad+\cos[(\omega_{k+l}-\omega_l)t-\phi_{k+l}+\phi_l-\omega_{k+l}\tau]\}\\ &=\tfrac{1}{2}C_{k+l}C_l\{(\cos\omega_l\tau+\cos\omega_{k+l}\tau)\cos(\omega_k t-\phi_{k+l}+\phi_l)\\ &\qquad-(\sin\omega_l\tau-\sin\omega_{k+l}\tau)\sin(\omega_k t-\phi_{k+l}+\phi_l)\}. \end{aligned} \tag{4}$$

The average power contributed by this particular set of terms is given by the expression

$$\begin{aligned} &\tfrac{1}{8}(C_{k+l}C_l)^2\{(\cos\omega_l\tau+\cos\omega_{k+l}\tau)^2+(\sin\omega_l\tau-\sin\omega_{k+l}\tau)^2\}\\ &\qquad=\tfrac{1}{4}(C_{k+l}C_l)^2\{1+\cos(\omega_l+\omega_{k+l})\tau\}. \end{aligned} \tag{5}$$

The corresponding terms for other values of l are random in phase relative to the above and to each other. Hence the total average power in the range f_k, $f_k+\Delta f$ due to the difference terms is obtained by summing Eq. (5) over all l.

$$\text{Av. power}=\frac{1}{4}\sum_{l=1}^{\infty} C_{k+l}^2 C_l^2(1+\cos\omega_{2l+k}\tau). \tag{6}$$

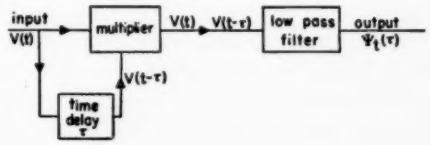


FIG. 3. The short-time autocorrelator.

The coefficients C_{k+l}^2 and C_l^2 may now be replaced by their equivalents $2w(f_{k+l})\Delta f$ and $2w(f_l)\Delta f$, respectively. If the frequency increment Δf becomes exceedingly small, the summation in l approaches an integral over all positive f and may be written in the following form.

$$\begin{aligned} \text{Av. power}\to\Delta f\int_0^{\infty} &w(f_k+f)w(f)\\ &\times[1+\cos 2\pi(2f+f_k)\tau]df. \end{aligned} \tag{7}$$

Analysis of the sum type modulation products proceeds in analogous manner. Here one begins with the pair of terms $m=k-l$, $n=l$, and $m=l$, $n=k-l$. The summation in l now extends only over the range $(1,k)$ and in doing so covers every term twice. Otherwise the computation is straightforward so that the power spectrum $W(f)$ of $V(t)V(t-\tau)$ can be written down without further difficulty.

$$\begin{aligned} W(f)&=\int_0^{\infty} w(x)w(f+x)[1+\cos 2\pi(2x+f)\tau]dx\\ &\quad+\frac{1}{2}\int_0^{f} w(x)w(f-x)[1+\cos 2\pi(2x-f)\tau]dx. \end{aligned} \tag{8}$$

If one extends the definition of $w(x)$ to negative frequencies in such a manner that $w(-x)\equiv w(x)$, one can easily establish that Eq. (8) may be converted to the more compact form

$$W(f)=\frac{1}{2}\int_{-\infty}^{\infty} dx w(x)w(f-x)[1+\cos 2\pi(2x-f)\tau]. \tag{9}$$

The power spectrum $W_0(f)$ after the low-pass filter is readily obtained from Eq. (9) since the filter is presumably a linear device with transfer function $F_0(2\pi jf)$.

$$W_0(f)=|F_0(2\pi jf)|^2 W(f). \tag{10}$$

As soon as $w(f)$ and $F_0(2\pi jf)$ are given, the analysis will be reduced to a purely computational problem.

THE FREQUENCY DISCRIMINATOR

A simple form of the standard frequency discriminator is shown in Fig. 4. F_1 and F_2 are two band pass filters with center frequencies f_1 and f_2 arranged symmetrically about the nominal mean frequency of the device. Square law detectors have been substituted for the more conventional linear rectifiers as a matter of mathematical convenience. This choice should not affect

the basic results to any appreciable degree. One may think of the system as operating by comparing the mean square voltage in the pass bands of the two filters. If F_1 and F_2 are made identical except for their center frequencies, the discriminator will indicate a mean frequency coinciding roughly with the center of mass of the incoming power spectrum. If the spectrum is symmetrical and unimodal, that quantity will, of course, be identical with the frequency of the peak.

The analysis now proceeds much as in the case of the short-time autocorrelator. Let the input voltage e_1 be represented again by the Fourier series of Eq. (2), and let the filter characteristics take the form

$$\left.\begin{aligned}\frac{e_2(\omega)}{e_1(\omega)}&=F_1(\omega)e^{i\psi_1(\omega)} \qquad (11)\\ \frac{e_3(\omega)}{e_1(\omega)}&=F_2(\omega)e^{i\psi_2(\omega)} \qquad (12)\end{aligned}\right\} F_1, F_2, \psi_1, \psi_2 \text{ real functions of } \omega.$$

It is an easy matter to write e_6 in terms of these quantities in form of sum and difference frequencies as in Eq. (3). The power spectrum $W_6(f)$ of e_6 is then again derived by separate examination of the sum and difference components. The details of this development are omitted here because it is algebraically tedious without offering any theoretical difficulties worthy of mention.

$$\begin{aligned}W_6(f)=2\int_0^\infty &\{F_1^2(f+x)F_1^2(x)+F_2^2(f+x)F_2^2(x)\\ &-2F_1(f+x)F_1(x)F_2(f+x)F_2(x)\\ &\times\cos[\psi_1(f+x)-\psi_1(x)-\psi_2(f+x)\\ &+\psi_2(x)]\}w(f+x)w(x)dx\\ +\int_0^f &\{F_1^2(f-x)F_1^2(x)+F_2^2(f-x)F_2^2(x)\\ &-2F_1(x)F_1(f-x)F_2(x)F_2(f-x)\\ &\times\cos[\psi_1(f-x)+\psi_1(x)-\psi_2(f-x)\\ &-\psi_2(x)]\}w(f-x)w(x)dx. \qquad (13)\end{aligned}$$

In complete analogy with Eq. (8), Eq. (13) can be converted to a more compact form if one assumes $w(-x)=w(x)$, $F_1(-x)=F_1(x)$, $F_2(-x)=F_2(x)$, $\psi_1(-x)=-\psi_1(x)$, $\psi_2(-x)=-\psi_2(x)$. The last four conditions are simply those of normal realizability in physical filters.

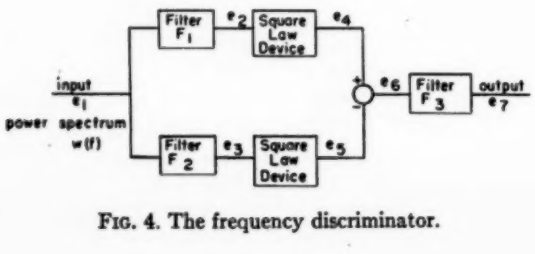


FIG. 4. The frequency discriminator.

$$\begin{aligned}W_6(f)=\int_{-\infty}^\infty &\{F_1^2(f-x)F_1^2(x)+F_2^2(f-x)F_2^2(x)\\ &-2F_1(f-x)F_1(x)F_2(f-x)F_2(x)\\ &\times\cos[\psi_1(f-x)+\psi_1(x)-\psi_2(f-x)\\ &-\psi_2(x)]\}w(f-x)w(x)dx. \qquad (14)\end{aligned}$$

The output power spectrum $W_7(f)$, finally, is obtained from the relation

$$W_7(f)=|F_0(2\pi jf)|^2W_6(f), \qquad (15)$$

where $F_0(2\pi jf)$ is the low-pass filter transfer function.

AUTOCORRELATOR AND DISCRIMINATOR PERFORMANCE IN THE ABSENCE OF WIDE BAND NOISE

Equations (9) and (13) will serve as the starting point for a discussion of system performance in response to an input consisting only of the narrow signal spectrum described in the introduction.

$$w(x)=\frac{\psi_0}{(2\pi)^{\frac{1}{2}}\sigma}\left\{\exp\left[-\frac{(x-f_0)^2}{2\sigma^2}\right]+\exp\left[-\frac{(x+f_0)^2}{2\sigma^2}\right]\right\}. \qquad (16)$$

Consider first the short-time autocorrelator. Substituting Eq. (16) into Eq. (9) and integrating, one obtains after some algebraic manipulation

$$\begin{aligned}W(f)=\frac{\psi_0^2}{4(\pi)^{\frac{1}{2}}\sigma}\Big\{&2[1+\exp(-4\pi^2\tau^2\sigma^2)\cos 4\pi f_0\tau]\\ &\times\exp\left(-\frac{f^2}{4\sigma^2}\right)+[1+\exp(-4\pi^2\tau^2\sigma^2)]\\ &\times\left[\exp\left(-\frac{[f-2f_0]^2}{4\sigma^2}\right)+\exp\left(-\frac{[f+2f_0]^2}{4\sigma^2}\right)\right]\Big\}. \qquad (17)\end{aligned}$$

The function $W(f)$ is sketched in Fig. 5. Any subsequent low-pass filter of narrow band width will effectively eliminate the double frequency portions of the spectrum, so that only the first term in Eq. (17) need be considered. That portion of $W(f)$ could also have been obtained by considering only the first term of Eq. (8) and using the input spectrum

$$w(x)=\frac{\psi_0}{(2\pi)^{\frac{1}{2}}\sigma}\exp\left[-\frac{(x-f_0)^2}{2\sigma^2}\right]. \qquad (18)$$

Because of the resultant reduction in complexity, that procedure will be employed in future calculations.[6]

The significant term of Eq. (17) has a form very important to the success of an autocorrelator instrumentation.

$$W(f)\cong\frac{\psi_0^2}{2(\pi)^{\frac{1}{2}}\sigma}[1+\exp(-4\pi^2\tau^2\sigma^2)\cos 4\pi f_0\tau]\times\exp\left(-\frac{f^2}{4\sigma^2}\right). \quad (19)$$

Considered as a function of τ the cosine term assumes the value

$$-\exp(-4\pi^2\tau^2\sigma^2)=-\exp\left[-\frac{\pi^2}{4}\frac{\sigma^2}{f_0^2}(2n+1)^2\right]$$

$$\text{at}\quad \tau=\frac{1}{4f_0}(2n+1),\ n=0,\ 1,\ 2,\ \cdots.$$

With $\sigma\ll f_0$, this value is very nearly equal to -1 for the lowest n. It follows that $W(f)$ and hence the rms output fluctuation exhibits a series of minima near the points $\tau=(1/4f_0)(2n+1)$. The first of these ($n=0$) is the sharpest and hence represents an optimum operating condition from this point of view.

The dc autocorrelator output is given by the expression

$$\psi(\tau)=\psi_0\exp(-2\pi^2\tau^2\sigma^2)\cos 2\pi f_0\tau. \quad (20)$$

The sensitivity S is therefore given by

$$S=\frac{\partial\psi(\tau,f_0)}{\partial f_0}=-2\pi\tau\psi_0\exp(-2\pi^2\tau^2\sigma^2)\sin 2\pi f_0\tau. \quad (21)$$

This quantity reaches a maximum near $\tau=(1/4f_0)\times(2n+1)$, $n=0,\ 1,\ 2\cdots$. The maximum for $\tau\cong 1/4f_0$ is the highest because of the exponential function in Eq. (21). Thus the rms fluctuation reaches a minimum essentially at the same point as the sensitivity attains its maximum value, so that the figure of merit μ_{11_0}/S^2 is clearly optimized at this point. It follows that the autocorrelator should be used as a null device operating about the first null of the true autocorrelation function $\psi(\tau)$. It is therefore well suited for operation in the circuit of Fig. 2. The time delay τ should be so adjusted that the relation $\tau=1/4f_0$ is readily satisfied when all loop components are operating within their linear ranges.

The autocorrelator figure of merit is now easily computed from Eqs. (10), (18), and (21). For simplicity in calculation a Gaussian low-pass filter will be used, although equivalent computations for physically realizable filter characteristics could be carried out without undue difficulty.

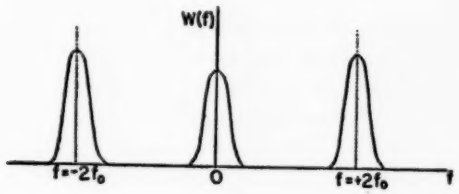


FIG. 5. Autocorrelator output spectrum.

$$|F_0(2\pi jf)|^2=\exp\left[-\frac{f^2}{2\sigma_0^2}\right],\quad \sigma_0\ll\sigma. \quad (22)$$

Then,

$$\mu_{11}=\int_0^\infty|F_0(2\pi jf)|^2W(f)df\cong\frac{\psi_0^2\sigma_0}{2\sqrt{2}\sigma}[1+\exp(-4\pi^2\tau^2\sigma^2)\cos 4\pi f_0\tau]. \quad (23)$$

Hence, the figure of merit at the operating null is given by the expression

$$\left.\frac{\mu_{11_0}}{S^2}\right|_{\tau=1/4f_0}=\frac{\sqrt{2}f_0^2\sigma_0\left[1-\exp\left(-\frac{\pi^2}{4}\frac{\sigma^2}{f_0^2}\right)\right]}{\sigma\pi^2\exp\left(-\frac{\pi^2}{4}\frac{\sigma^2}{f_0^2}\right)}. \quad (24)$$

Equation (24) has dimensions of [frequency]^2. This is reasonable because it represents the mean square fluctuation of the instantaneous measurement of f_0 about its average value. An absolute figure of merit, the fractional error in short-time measurement, is obtained by normalizing with respect to f_0^2. Dividing Eq. (24) by this factor and using the numerical ratio $\sigma/f_0=0.15$, one obtains the result

$$\frac{\mu_{11_0}}{f_0^2S^2}=0.00825\frac{\sigma_0}{\sigma}. \quad (25)$$

This figure must now be compared with the corresponding quantity for the discriminator instrumentation.

Before the discriminator output spectrum can be evaluated from Eq. (13) or (14), the two filter characteristics must be specified. Since the exact nature of the amplitude functions $F_1(x)$ and $F_2(x)$ should not influence such design considerations as filter band width and separation very critically, a Gaussian amplitude characteristic will be assumed.

$$F_1^2(x)=\exp\left[-\frac{(x-f_1)^2}{2\sigma_f^2}\right]$$

$$F_2^2(x)=\exp\left[-\frac{(x-f_2)^2}{2\sigma_f^2}\right]. \quad (26)$$

[6] A slight discrepancy between the two methods arises from the fact that Eq. (18) does not vanish identically for $x<0$. However, with $\sigma\ll f_0$, one may always replace an integral of the form

$$\int_0^\infty\exp\left[-\frac{(x-f_0)^2}{2\sigma^2}\right]dx \quad\text{by}\quad \int_{-\infty}^\infty\exp\left[-\frac{(x-f_0)^2}{2\sigma^2}\right]dx$$

without introducing any significant error.

The question of a reasonable phase function remains to be decided. It is not possible to compute a consistent phase function from Bode's integral because the Gaussian is not physically realizable and the resulting integral diverges. The choice must therefore be somewhat arbitrary. A brief inspection of Eq. (13) reveals that this difficulty is not as serious as it might appear, for the phase function enters into the significant first term only in form of the difference $\psi_1(f+x)-\psi_1(x)-[\psi_2(f+x)-\psi_2(x)]$. With a low-pass filter of less than 1 cps band width, all f of interest are exceedingly small. The discriminator filters, with band widths of the order of 100 cps or more, cannot exhibit appreciable phase shift over so short an interval, and if the two filters are similar, even these small phase shifts tend to cancel. An approximation for ψ_1 and ψ_2 using only the first terms of a power series expansion should therefore be more than ample.

$$\psi_1(x)=-a(x-f_1),\quad \psi_2(x)=-a(x-f_2),\quad a>0. \tag{27}$$

Then the argument of the cosine term in Eq. (13) is zero and the equation is simplified considerably. Equations (26) and (27) may now be substituted into the first term of Eq. (13). In the low-frequency range the spectrum then assumes the form

$$W_6(f)=\frac{\psi_0^2\rho}{\sqrt{\pi}\sigma^2}\left\{\exp\left[-\frac{(f_0-f_1)^2}{\sigma^2+\sigma_f^2}\right]+\exp\left[-\frac{(f_0-f_2)^2}{\sigma^2+\sigma_f^2}\right]-2\exp\left[-\frac{(f_1-f_2)^2}{4\sigma_f^2}\right]\exp\left[-\frac{(f_m-f_0)^2}{\sigma^2+\sigma_f^2}\right]\right\}\times\left\{\exp\left[-\frac{f^2}{4\rho^2}\right]\right\}, \tag{28}$$

where

$$\rho\equiv\frac{\sigma\sigma_f}{(\sigma^2+\sigma_f^2)^{\frac{1}{2}}},\quad f_m\equiv\frac{f_1+f_2}{2}.$$

The power spectrum after the low-pass filter is obtained from Eq. (15), Eq. (22) being used to represent the low-pass filter characteristic. The output variance follows immediately from integration of $W_7(f)$.

$$\text{Variance}=\int_0^\infty W_7(f)df=\frac{\psi_0^2\sigma_0\rho}{\sqrt{2}\sigma^2}\times\left\{\exp\left[-\frac{(f_0-f_1)^2}{\sigma^2+\sigma_f^2}\right]+\exp\left[-\frac{(f_0-f_2)^2}{\sigma^2+\sigma_f^2}\right]-2\exp\left[-\frac{(f_1-f_2)^2}{4\sigma_f^2}\right]\exp\left[-\frac{(f_m-f_0)^2}{\sigma^2+\sigma_f^2}\right]\right\}, \tag{29}$$

where use has been made of the fact that $\sigma_0\ll\rho$. The dc output $\bar{e}_7$ of the discriminator is given by the expression

$$\bar{e}_7=\int_0^\infty[F_1^2(f)-F_2^2(f)]w(f)df=\frac{\psi_0\rho}{\sigma}\left\{\exp\left[-\frac{(f_0-f_1)^2}{2(\sigma^2+\sigma_f^2)}\right]-\exp\left[-\frac{(f_0-f_2)^2}{2(\sigma^2+\sigma_f^2)}\right]\right\}. \tag{30}$$

The sensitivity is obtained by differentiating Eq. (30) with respect to f_0 and the figure of merit for the operating null can now be calculated without difficulty. After some simplification it takes the form

$$\left.\frac{\mu_{11_0}}{S^2}\right|_{f_0=f_m}=\frac{\sqrt{2}\sigma_0(\sigma^2+\sigma_f^2)^{5/2}}{\sigma\sigma_f(f_1-f_2)^2}\times\left\{1-\exp\left[-\frac{(f_1-f_2)^2}{4\frac{\sigma_f^2}{\sigma^2}(\sigma^2+\sigma_f^2)}\right]\right\}. \tag{31}$$

An attempt to minimize Eq. (31) with respect to the filter separation (f_1-f_2) leads to the result $(f_1-f_2)\to\infty$. In physical terms this implies that the sensitivity falls off less rapidly than the output fluctuation so that the theoretical optimum performance would be obtained with the widest possible filter separation. Since some noise is always present in fact, such an adjustment would be entirely unrealistic. No complete evaluation of the frequency discriminator can therefore be made before the general case, including noise as well as signal, has been considered. This is done in the following section. For the time being, in the interest of a preliminary comparison with the autocorrelator, an arbitrary figure, justified only by the argument of the subsequent section, will be assumed. Using $(f_1-f_2)^2=6\sigma^2$, $\sigma=0.15\,f_0$, and normalizing Eq. (31) with respect to f_0^2 one obtains the figure of merit

$$\left.\frac{\mu_{11_0}}{f_0^2S^2}\right|_{f_0=f_m}=0.0053\frac{\sigma_0}{\sigma}\frac{\left(1+\frac{\sigma_f^2}{\sigma^2}\right)^{5/2}}{\frac{\sigma_f}{\sigma}}\times\left\{1-\exp\left[-\frac{1.5}{\frac{\sigma_f^2}{\sigma^2}\left(1+\frac{\sigma_f^2}{\sigma^2}\right)}\right]\right\}. \tag{32}$$

A plot of Eq. (32) is given in the lowest curve of Fig. 10. The minimum occurs at $\sigma/\sigma_f=0$ and has the value

$$\frac{\mu_{11_0}}{f_0^2S^2}=0.0079\frac{\sigma_0}{\sigma}. \tag{33}$$

This compares quite closely with the corresponding figure [Eq. (25)] for the autocorrelator. It should be pointed out, however, that for all practical purposes in the autocorrelator the condition of best figure of merit coincides with that of maximum sensitivity, while in the discriminator the sensitivity approaches zero as the figure of merit reaches its minimum. This argument, apparently in favor of the autocorrelator, is of course strictly preliminary and subject to change when inputs with finite signal-to-noise ratio are considered.

AUTOCORRELATOR AND DISCRIMINATOR PERFORMANCE IN THE PRESENCE OF WIDE-BAND NOISE

The present section discusses the most general case, where the system input consists of wide-band noise superimposed on the narrow-band signal considered thus far. Unfortunately the complexity of mathematical manipulation is such as to make it impractical to incorporate detailed derivations in this paper. After a statement of the initial equations the result will therefore be given immediately in the most general form available. Thereafter a particular ratio of signal band-width to center frequency as well as a narrow low-pass filter will be assumed in order to obtain a meaningful comparison under typical operating conditions. Since the difficulties alluded to are entirely due to algebraic complexity, the reader should have no trouble in modifying the general results to fit other particular situations.

If signal and noise are independent, the input power spectrum can be represented by the expression

$$w(f)=\frac{\psi_0}{(2\pi)^{\frac{1}{2}}\sigma}\exp\left[-\frac{(f-f_0)^2}{2\sigma^2}\right]+\psi_n. \tag{34}$$

Note that ψ_n is the noise power per unit frequency whereas ψ_0 is the total signal power. The total noise power represented by Eq. (34) is infinite, hence the autocorrelator must be equipped with a frequency selective filter if the output fluctuation is to remain within tolerable limits. The circuit of Fig. 3 is therefore modified and now assumes the form of Fig. 6. The band-pass filter is tuned to the frequency at which the delay line introduces 90° phase shift and is assumed to have the Gaussian transfer function

$$|H(f)|^2=\exp\left[-\frac{(f-f_1)^2}{2\sigma_f^2}\right]. \tag{35}$$

It is immediately evident that the voltage V_1 in Fig. 6 is normally distributed in amplitude and has the power spectrum

$$w_1(f)=w(f)\exp\left[-\frac{(f-f_1)^2}{2\sigma_f^2}\right]. \tag{36}$$

If Eq. (36) is interpreted as the input power spectrum, the first term of Eq. (8) may be used without

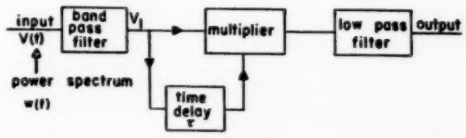


FIG. 6. Short-time auto-correlator with noise limiting filter.

modification to represent the low-frequency portion $W(f)$ of the spectrum ahead of the low-pass filter. Formal substitution of Eqs. (36) and (34) into the first term of Eq. (8) leads to the result

$$\begin{aligned}W(f)=&\frac{\psi_0^2}{2(\pi)^{\frac{1}{2}}}\frac{\sigma_f}{\sigma(\sigma^2+\sigma_f^2)^{\frac{1}{2}}}\exp\left[-\frac{(f_0-f_1)^2}{\sigma^2+\sigma_f^2}\right]\\&\times\{1+\exp[-4\pi^2\tau^2(\sigma^2*\sigma_f^2)]\cos 4\pi(f_0*f_1)\tau\}\\&+\psi_0\psi_n\frac{\left(\sigma^2*\frac{\sigma_f^2}{2}\right)^{\frac{1}{2}}}{\sigma}\exp\left[-\frac{(f_1-f_0)^2}{2(\sigma^2+\sigma_f^2)}\right]\\&\times\left\{\exp\left[-\frac{(f_1+f-f_0*f_1)^2}{2(\sigma^2*\sigma_f^2+\sigma_f^2)}\right]\right.\\&\times\left[1+\exp\left(-8\pi^2\tau^2\sigma^2*\frac{\sigma_f^2}{2}\right)\right.\\&\left.\times\cos\left(2\pi f\tau\frac{\sigma^2*\sigma_f^2-\sigma_f^2}{\sigma^2*\sigma_f^2+\sigma_f^2}+4\pi\tau(f_0*f_1)*f_1\right)\right]\\&+\exp\left[-\frac{(f_1-f-f_0*f_1)^2}{2(\sigma^2*\sigma_f^2+\sigma_f^2)}\right]\\&\times\left[1+\exp\left(-8\pi^2\tau^2\sigma^2*\frac{\sigma_f^2}{2}\right)\right.\\&\left.\left.\times\cos\left(2\pi f\tau\frac{\sigma_f^2-\sigma^2*\sigma_f^2}{\sigma^2*\sigma_f^2+\sigma_f^2}+4\pi\tau(f_0*f_1)*f_1\right)\right]\right\}\\&+(\pi)^{\frac{1}{2}}\sigma_f\psi_n^2\{1+\exp[-4\pi^2\tau^2\sigma_f^2]\\&\times\cos 4\pi f_1\tau\}\exp\left[-\frac{f^2}{4\sigma_f^2}\right],\end{aligned} \tag{37}$$

where

$$f_a*f_b\equiv\frac{f_a\sigma_b^2+f_b\sigma_a^2}{\sigma_a^2+\sigma_b^2},\quad \sigma_a^2*\sigma_b^2\equiv\frac{\sigma_a^2\sigma_b^2}{\sigma_a^2+\sigma_b^2}.$$

The computation of output variance now requires only multiplication of Eq. (37) by the square of the low-pass filter characteristic (Eq. (22)) and integration over all positive f. This process becomes elementary once one recognizes that the two terms in $\psi_0\psi_n$ are symmetrical with respect to $f=0$ and can be treated as a unit. Under assumption of a narrow low-pass filter

($\sigma_0 \ll \sigma$) the result is

$$\text{variance} = \psi_0^2 \frac{\sigma_f \sigma_0}{2\sqrt{2}\sigma(\sigma^2+\sigma_f^2)^{\frac{1}{2}}} \exp\left[-\frac{(f_0-f_1)^2}{\sigma^2+\sigma_f^2}\right]$$
$$\times\left\{1+\exp\left[-4\pi^2\tau^2\frac{\sigma^2\sigma_f^2}{\sigma^2+\sigma_f^2}\right]\cos 4\pi\tau f_0{}^* f_1\right\}$$
$$+\psi_0(\psi_n\sigma)\frac{(2\pi)^{\frac{1}{2}}\sigma_f\sigma_0}{\sigma(2\sigma^2+\sigma_f^2)^{\frac{1}{2}}}\exp\left[-\frac{(f_1-f_0)^2}{2(\sigma^2+\sigma_f^2)}\right]$$
$$\times\exp\left[-\frac{(f_1-f_0{}^* f_1)^2}{2(\sigma^2 * \sigma_f^2+\sigma_f^2)}\right]$$
$$\times\left\{1+\exp\left[-8\pi^2\tau^2\frac{\sigma^2\sigma_f^2}{\sigma_f^2+2\sigma^2}\right]\right.$$
$$\left.\times\cos 4\pi\tau(f_0{}^* f_1)^* f_1\right\}+(\psi_n\sigma)^2\frac{\pi}{\sqrt{2}}\frac{\sigma_f\sigma_0}{\sigma^2}$$
$$\times\{1+\exp[-4\pi^2\tau^2\sigma_f{}^2]\cos 4\pi f_1\tau\}. \quad (38)$$

The sensitivity remains to be computed. Since wide-band noise does not contribute any dc output, the sensitivity can be calculated from the signal autocorrelation alone.

$$\frac{\partial\psi(\tau,f_0)}{\partial f_0}=\psi_0\frac{(\sigma_f^2 * \sigma^2)^{\frac{1}{2}}}{\sigma}\exp[-2\pi^2\tau^2\sigma^2 * \sigma_f{}^2]$$
$$\times\left\{-2\pi\tau\frac{\sigma_f^2}{\sigma^2+\sigma_f^2}\exp\left[-\frac{(f_1-f_0)^2}{2(\sigma^2+\sigma_f^2)}\right]\right.$$
$$\times\sin 2\pi(f_0{}^* f_1)\tau+\frac{f_1-f_0}{\sigma^2+\sigma_f^2}$$
$$\left.\times\exp\left[-\frac{(f_1-f_0)^2}{2(\sigma^2+\sigma_f^2)}\right]\cos 2\pi(f_0{}^* f_1)\tau\right\}. \quad (39)$$

After Eqs. (38) and (39) have been evaluated at the operating null $\tau=1/4f_0=1/4f_1$ the usual figure of merit can be computed without difficulty. Normalizing with respect to f_0^2 and assuming the numerical relation $\sigma=0.15f_0$ one obtains the relation

$$\frac{\text{variance}}{f_0^2(\text{sensitivity})^2}=\frac{\mu_{11_0}}{f_0^2 S^2}$$
$$=\frac{\sigma_0}{\sigma}\frac{\frac{1}{2\sqrt{2}}\frac{1}{(1+x)^{\frac{1}{2}}}\left\{1-\exp\left[-\frac{0.0555}{1+x}\right]\right\}+\left(\frac{\psi_N\sigma}{\psi_0}\right)\frac{(2\pi)^{\frac{1}{2}}}{(1+2x)^{\frac{1}{2}}}\left\{1-\exp\left[-\frac{0.111}{1+2x}\right]\right\}+\left(\frac{\psi_N\sigma}{\psi_0}\right)^2\frac{\pi}{\sqrt{2}}\frac{1}{x}\left\{1-\exp\left[-\frac{0.0555}{x}\right]\right\}}{2.46\frac{1}{(1+x)^3}\exp\left[-\frac{0.0555}{1+x}\right]}, \quad (40)$$

where

$$x\equiv\frac{\sigma^2}{\sigma_f{}^2}.$$

Note that $\psi_N\sigma/\psi_0$ is the effective noise-to-signal ratio since $\psi_N\sigma$ is the total noise power present in a band equivalent to that of the signal. The definition, of course, involves a certain degree of arbitrariness because the signal power is distributed in a Gaussian manner whereas the noise spectrum is flat.

Equation (40) may be used to compute optimum values of x for various noise levels. The process is a purely numerical one and the results are plotted in Figs. 7, 8, and 9. Figure 7 shows plots of the figure of merit *versus* filter band width with the signal-to-noise ratio as a parameter. It will be observed that the minima shift to narrower filter band widths as the signal-to-noise ratio decreases. The required change in band width for optimum performance is not very drastic, however, if one excludes from consideration noise levels below 10 percent of the signal. Thus between $\psi_N\sigma/\psi_0=0.1$ and $\psi_N\sigma/\psi_0=1.0$ the ratio σ/σ_f should change only by a factor $(3)^{\frac{1}{2}}$. This information is readily obtained from

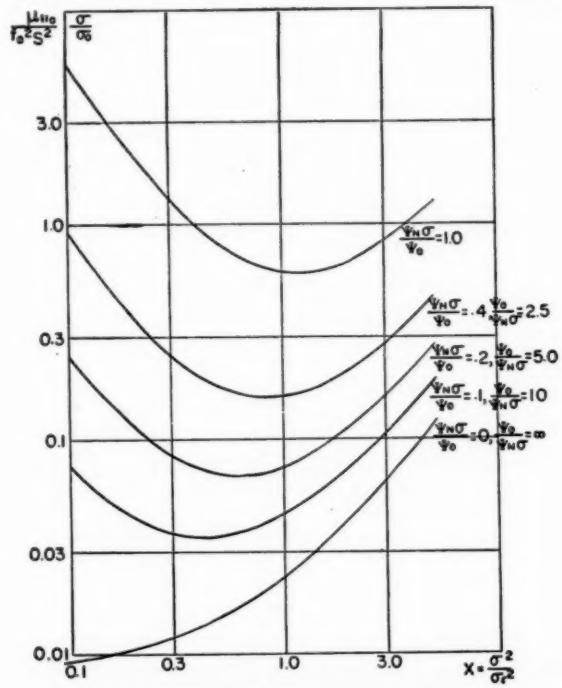


FIG. 7. Autocorrelator steady-state performance.

Fig. 8, which shows the proper filter adjustment as a function of noise-to-signal ratio. The use of linear scales emphasizes the relatively slight filter adjustments in the range normally of interest.

In practice it will not be possible, of course, to adjust the filter band width for an optimum at every signal-to-noise ratio. One will probably choose the filter so as to optimize performance at the highest expected noise level since the tracking difficulties are most serious under that condition. The reduction in accuracy resulting from this procedure is indicated in Fig. 9. Evidently the loss is serious only at very low noise levels, in a range unlikely to be of great practical importance.

Figure 9 also suggests a rapid increase in tracking difficulty as the signal-to-noise ratio approaches unity. The fluctuation rises almost with the second power of $\psi_N\sigma/\psi_0$ in that neighborhood.

The low frequency output spectrum of the frequency discriminator can be obtained directly from the first term of Eq. (13). Equation (34) is now used to represent the input spectrum, otherwise the procedure does not differ from that for the noise-free case. After considerable mathematical manipulation the result assumes the form

$$
\begin{aligned}
W_6(f)=&\frac{\psi_0^2(\sigma^2*\sigma_f^2)^{\frac{1}{2}}}{(\pi)^{\frac{1}{2}}\sigma^2}\exp\left[-\frac{f^2}{4\sigma^2*\sigma_f^2}\right]\\
&\times\left\{\exp\left[-\frac{(f_1-f_0)^2}{\sigma^2+\sigma_f^2}\right]+\exp\left[-\frac{(f_2-f_0)^2}{\sigma^2+\sigma_f^2}\right]\right.\\
&\left.-2\exp\left[-\frac{(f_1-f_2)^2}{4\sigma_f^2}\right]\exp\left[-\frac{(f_m-f_0)^2}{\sigma_f^2+\sigma^2}\right]\right\}\\
&+2\psi_0\psi_N\frac{(\sigma^2*\sigma_f^2*\sigma_f^2)^{\frac{1}{2}}}{\sigma}\left\{\exp\left[-\frac{(f_1-f_0)^2}{2(\sigma^2+\sigma_f^2)}\right]\right.\\
&\times\left(\exp\left[-\frac{(f+f_0{}^*f_1-f_1)^2}{2(\sigma_f^2+\sigma^2*\sigma_f^2)}\right]\right.\\
&\left.+\exp\left[-\frac{(f+f_1-f_0{}^*f_0)^2}{2(\sigma_f^2+\sigma^2*\sigma_f^2)}\right]\right)\\
&+\exp\left[-\frac{(f_2-f_0)^2}{2(\sigma^2+\sigma_f^2)}\right]\left(\exp\left[-\frac{(f+f_0{}^*f_2-f_2)^2}{2(\sigma_f^2+\sigma^2*\sigma_f^2)}\right]\right.\\
&\left.+\exp\left[-\frac{(f+f_2-f_0{}^*f_2)^2}{2(\sigma_f^2+\sigma^2*\sigma_f^2)}\right]\right)-2\exp\left[-\frac{(f_1-f_2)^2}{4\sigma_f^2}\right]\\
&\times\exp\left[-\frac{(f_m-f_0)^2}{2(\sigma_f^2+\sigma^2)}\right]\left(\exp\left[-\frac{(f+f_m{}^*f_0-f_m)^2}{2(\sigma_f^2+\sigma^2*\sigma_f^2)}\right]\right.\\
&\left.\left.+\exp\left[-\frac{(f+f_m-f_m{}^*f_0)^2}{2(\sigma_f^2+\sigma^2*\sigma_f^2)}\right]\right)\right\}+4(\pi)^{\frac{1}{2}}\psi_N{}^2\sigma_f\\
&\times\exp\left[-\frac{f^2}{4\sigma_f^2}\right]\left\{1-\exp\left[-\frac{(f_1-f_2)^2}{4\sigma_f^2}\right]\right\},
\end{aligned}
\tag{41}
$$

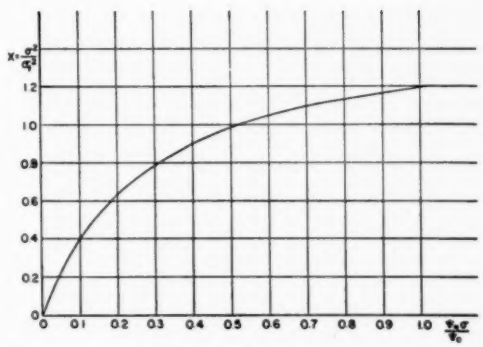


FIG. 8. Autocorrelator filter adjustment for optimum performance.

where

$$f_m\equiv\frac{f_1+f_2}{2}.$$

Assuming again a narrow Gaussian low-pass filter [Eq. (22)] and integrating over all positive f, one obtains the desired output variance

$$
\begin{aligned}
\text{variance}=&\frac{\sqrt{2}(\sigma^2*\sigma_f^2)^{\frac{1}{2}}\sigma_0}{\sigma^2}\left\{\exp\left[-\frac{(f_1-f_2)^2}{4(\sigma^2+\sigma_f^2)}\right]\right.\\
&\left.-\exp\left[-\frac{(f_1-f_2)^2}{4\sigma_f^2}\right]\right\}\psi_0{}^2+\frac{4(2\pi)^{\frac{1}{2}}\left(\sigma^2*\frac{\sigma_f^2}{2}\right)^{\frac{1}{2}}\sigma_0}{\sigma}\\
&\times\left\{\exp\left[-\frac{(f_1-f_2)^2}{8\frac{(\sigma_f^2+\sigma^2)}{\sigma_f^2}(\sigma_f^2+2\sigma^2)}\right]\right.\\
&\left.\times\exp\left[-\frac{(f_1-f_2)^2}{8(\sigma^2+\sigma_f^2)}\right]-\exp\left[-\frac{(f_1-f_2)^2}{4\sigma_f^2}\right]\right\}\psi_0\psi_N\\
&+2\sqrt{2}\pi\sigma_f\sigma_0\left\{1-\exp\left[-\frac{(f_1-f_2)^2}{4\sigma_f^2}\right]\right\}\psi_N{}^2.
\end{aligned}
\tag{42}
$$

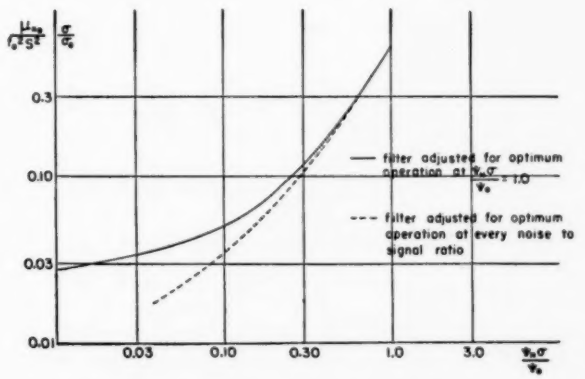


FIG. 9. Dependence of autocorrelator performance on noise-to-signal ratio.

The sensitivity to center frequency changes is clearly unaffected by the presence of wide-band noise so that the derivative of Eq. (30) remains applicable. The usual performance criterion variance/S^2 is now obtained by dividing the square of this sensitivity into (42) and normalizing with respect to f_0^2.

$$\frac{\mu_{11_0}}{f_0^2 S^2} = 1.41 \frac{\sigma_0}{\sigma} \frac{\sigma^2}{f_0^2} \frac{(\sigma^2+\sigma_f^2)^{5/2}}{\sigma^2\sigma_f(f_1-f_2)^2}$$
$$\times \left\{ \left[\left[1-\exp\left[-\frac{(f_1-f_2)^2}{4\sigma_f^2\left(1+\frac{\sigma_f^2}{\sigma^2}\right)} \right] \right] \right.$$
$$+ \frac{7.08}{\left[1+\frac{1}{1+\frac{\sigma_f^2}{\sigma^2}} \right]^{\frac{1}{2}}} \exp\left[\frac{(f_1-f_2)^2}{8(\sigma^2+\sigma_f^2)\left(1+\frac{\sigma_f^2}{2\sigma^2}\right)} \right]$$
$$\times \left[1-\exp\left[-\frac{(f_1-f_2)^2}{4\sigma_f^2\left(1+\frac{\sigma_f^2}{2\sigma^2}\right)} \right] \right] \left. \frac{\psi_N\sigma}{\psi_0} \right.$$
$$+ 6.28\left(1+\frac{\sigma_f^2}{\sigma^2}\right)^{\frac{1}{2}} \exp\left[\frac{(f_1-f_2)^2}{4(\sigma^2+\sigma_f^2)} \right]$$
$$\left. \times \left(1-\exp\left[-\frac{(f_1-f_2)^2}{4\sigma_f^2} \right] \right) \left(\frac{\psi_N\sigma}{\psi_0} \right)^2 \right\}. \qquad (43)$$

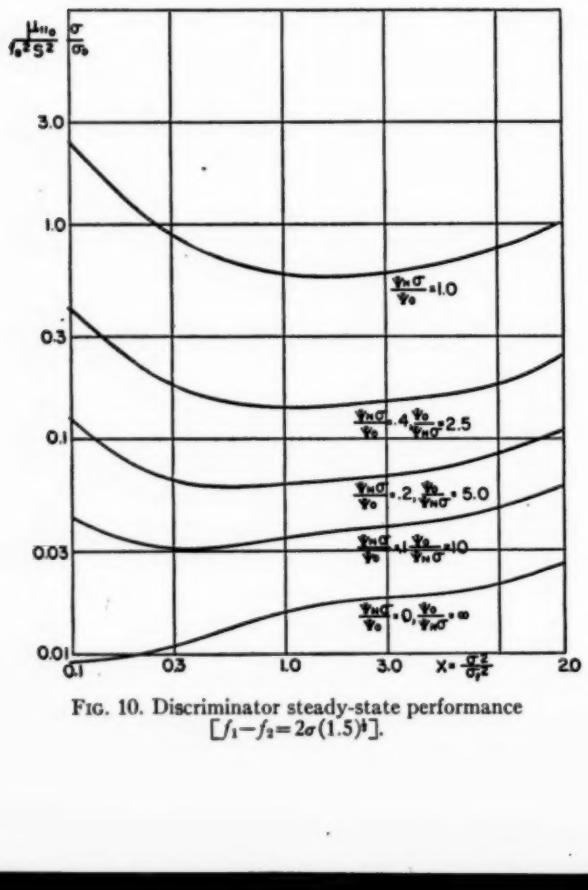


FIG. 10. Discriminator steady-state performance $[f_1-f_2=2\sigma(1.5)^{\frac{1}{2}}]$.

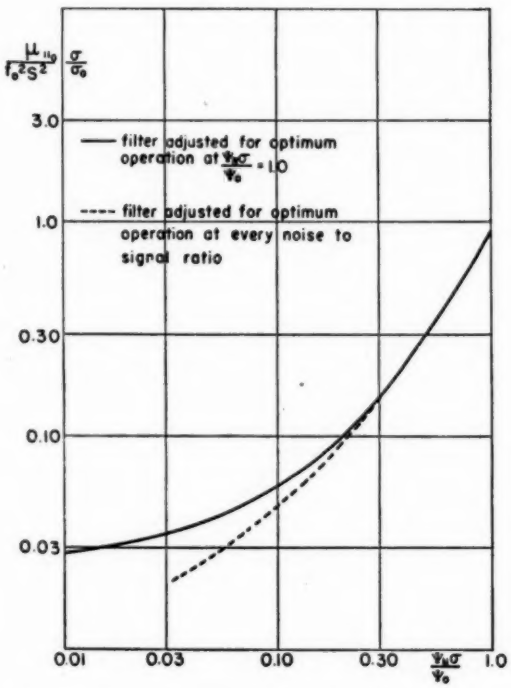


FIG. 11. Dependence of discriminator performance on noise-to-signal ratio.

Equation (43) is analogous to (40) in the autocorrelator discussion. However, its interpretation is more difficult because two parameters, the filter band width σ_f and separation (f_1-f_2), may be varied independently. Thus a set of curves is required to describe completely the performance at any one signal-to-noise ratio. The problem is simplified considerably by the practical consideration that all filter adjustments will have to be permanent. Performance can thus be optimized for any preassigned operating condition, such as the poorest signal-to-noise ratio expected, but once this choice has been made a decline in performance relative to the ideal value must be tolerated elsewhere. Selection of the proper operating point for this adjustment is a practical matter, determined largely by the anticipated duty cycle and disturbances. In the absence of quantitative information one tends to optimize performance under the worst possible conditions. This procedure will be adopted here and a 1:1 signal-to-noise ratio will be assumed.[7]

Even with this simplification, the selection of parameters from Eq. (43) remains basically a numerical process, for differentiation with respect to either σ_f or (f_1-f_2) does not lead to an expression readily interpreted in algebraic form. Rough calculations indicate

[7] One may easily conceive of a situation where this technique would not lead to the best result. If, for instance, the poorest signal-to-noise ratio were to arise only for rare and brief intervals, one would probably gain by tolerating large errors during these periods in return for better performance at other times. In such a case one could always return to Eq. (43) in an effort to strike the most advantageous compromise by numerical computation, a tedious but straightforward process.

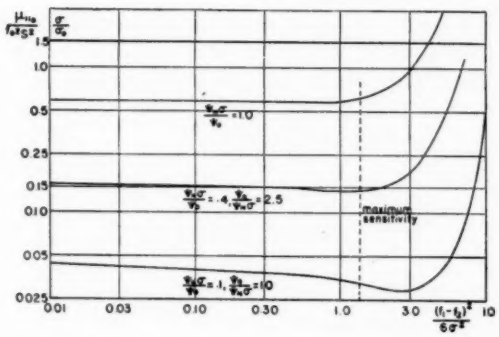


FIG. 12. Effect of filter separation on discriminator performance ($\sigma_f=\sigma$).

that the filter separation should approach the value providing maximum sensitivity as the noise level becomes high. This condition is reached when the following relation is satisfied

$$(f_1-f_2)=2(\sigma^2+\sigma_f^2)^{\frac{1}{2}}. \tag{44}$$

At lower noise levels the optimum spacing is somewhat larger, but the increase does not become very significant until the signal-to-noise ratio exceeds 10. Thus one may safely conclude that the adjustment should be close to the figure demanded by Eq. (44). Preliminary calculations also indicate that the filter band width for high noise levels should be approximately $\sigma/\sqrt{2}$.[8] With these considerations in mind, the filter separation, the less critical of the two independent variables, was fixed at

$$(f_1-f_2)=2\sigma(1.5)^{\frac{1}{2}}. \tag{45}$$

Substitution of (45) into (43) leads to a function of the single variable σ_f which may now be evaluated for various signal-to-noise ratios. The result is shown in Fig. 10. The independent variable $x=\sigma^2/\sigma_f^2$ has been used to facilitate comparison with the autocorrelator (Fig. 7).

Inspection of Figs. 7 and 10 leads to two conclusions:

(1) If both systems are adjusted for optimum operation at a given signal-to-noise ratio, their performance under that condition will be almost entirely equivalent.

(2) The autocorrelator is more sensitive to adjustment than the frequency discriminator. The latter has characteristics essentially flat over a 10:1 range of x while the autocorrelator characteristics exhibit considerable variation over a similar range. If the system is subjected to operating conditions other than those assumed for adjustment purposes, the discriminator is therefore preferable. This effect may become quite pronounced, particularly if a change in the standard deviation σ of the input power spectrum shifts the operating point far from its nominal value. Comparison of Figs. 7 and 10 reveals that a 3:1 or 4:1 advantage in favor of the discriminator may easily arise in such a situation.

[8] Filter band width for maximum dc sensitivity is $\sigma_f=\sigma$.

Figure 11, like its counterpart 9, indicates the reduction in discriminator accuracy resulting from the adjustment of parameters at a single assumed operating condition, here a 1:1 signal-to-noise ratio. As suggested in the above discussion, the loss is almost negligible throughout the region of practical interest.

Finally in Fig. 12, an attempt is made to investigate the dependence of Fig. 10 on filter separation. It will be recalled that the separation was chosen on the basis of rough calculations for 1:1 signal-to-noise ratio. Figure 12 justifies the selection, for a reduction in spacing at that noise level leads only to lower sensitivity without reduction in the figure of merit whereas an increase results in immediate and marked deterioration. There is no pronounced minimum in this curve in contrast to plots for higher signal-to-noise ratios. As pointed out earlier, some improvement can be secured by increasing filter separation as the noise level diminishes. Figure 12 shows, however, that the benefit obtainable by this procedure is quite limited even for signal-to-noise ratios as high as 10:1, while the accompanying loss at higher noise levels is disproportionately larger. Operation near the point of maximum sensitivity therefore appears almost imperative unless the upper limit of signal-to-noise ratio is known accurately.

ACKNOWLEDGMENT

The research described in this paper was done under contract to General Precision Laboratory,[9] Pleasantville, New York. The authors particularly wish to acknowledge the help of Dr. F. B. Berger of General Precision Laboratory who contributed many valuable suggestions in the course of this work.

APPENDIX

Derivation of the Figure of Merit

Consider a particular instrumentation such as the discriminator of Fig. 4. Its average output $\bar{e}_7$ is shown in Fig. 13, a curve which is approximately linear near $f_0=f_m$, the operating point in a null seeking system. If the discriminator input is a spectrum (Fig. 1), the instantaneous output $e_7(f_0,t)$ will fluctuate about zero.

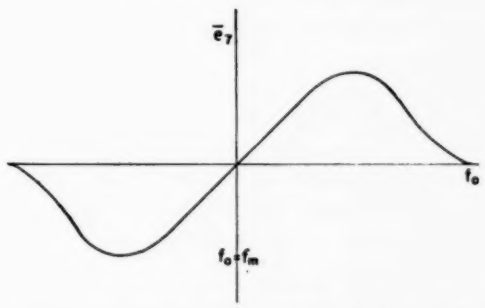


FIG. 13. Average discriminator output as a function of f_0.

[9] The work was sponsored indirectly by the United States Air Force.

However, for each t there will be a value of f_0 which satisfies $e_7(f_0,t)=0$ and it appears reasonable to speak of that value as the instantaneous measurement of f_0. The quality of performance is then determined by the probability distribution of the instantaneous measurement; one may speak of good performance if the distribution is narrow and poor performance if it is wide. The calculation of this distribution will now be carried out in fairly general terms.

It will be assumed that the narrow-band input spectrum is sufficiently narrow to be well defined but that its width is an appreciable percentage of f_0 (15 percent will be used in numerical calculation).[10] If high accuracies are to be obtained (better than 1 percent), the data must be smoothed for a rather long time, an operation demanding that the pass band of the low-pass filter must be very narrow compared to the input spectrum width. Under these circumstances Rice[11] and others have demonstrated that the filter output assumes an essentially normal probability distribution, a circumstance which, under certain reasonable assumptions, reduces the determination of the probability density for the instantaneous null to a fairly straightforward computational problem.

Let the frequency sensitive element have an output $G(x)$ which is normally distributed and whose long-time average has a null at $x=x_0$. What is the probability density of the coordinate x of an instantaneous null of $G(x)$ near x_0?

The computation is based on the following theorem due to Rice.[12]

If $y=F(a_1,a_2\cdots;x)$ a function of x with the a's random variables, the probability of y having a zero with positive slope in the interval $(x,x+dx)$ is given by the expression

$$dx\int_0^\infty \eta p(0,\eta;x)d\eta, \tag{46}$$

where $p(\xi,\eta;x)$ is the two-dimensional probability density function of the variables $\xi\equiv F(a_1,a_2\cdots;x)$ and $\eta\equiv\partial F/\partial x$. For the problem at hand, ξ will be identified with $G(x)$, η with $\partial G/\partial x$, and the a's will be random parameters such as the random phase angles in Eq. (2). Under the assumptions made $p(\xi,\eta;x)$ is normal in two dimensions:

$$p(\xi,\eta;x)=\frac{1}{2\pi|M|^{\frac{1}{2}}}\exp\left[-\frac{\mu_{11}}{2|M|}(\eta-\bar\eta)^2 - \frac{\mu_{22}}{2|M|}(\xi-\bar\xi)^2+\frac{2\mu_{12}}{2|M|}(\xi-\bar\xi)(\eta-\bar\eta)\right], \tag{47}$$

where $|M|\equiv\mu_{11}\mu_{22}-\mu_{12}^2$, $\mu_{11}\equiv\langle\xi^2\rangle_{Av}-\bar\xi^2$, $\mu_{22}=\langle\eta^2\rangle_{Av}-\bar\eta^2$, $\mu_{12}=\langle\xi\eta\rangle_{Av}-(\bar\xi)(\bar\eta)$. Substituting Eq. (47) into Eq. (46) and integrating, one obtains the approximate expression for the probability density $p(x)$ of a zero with positive slope at x

$$p(x)\cong\frac{1}{(2\pi\mu_{11})^{\frac{1}{2}}}\left(\bar\eta-\frac{\mu_{12}}{\mu_{11}}\bar\xi\right)\exp\left[-\frac{\bar\xi^2}{2\mu_{11}}\right]. \tag{48}$$

The approximation consists in extension of the range of integration to the interval $(-\infty,\infty)$, justified by the observation that the one-dimensional marginal distribution in η will be narrow if a sufficiently large low-pass filter is used, or in other words that the occurrence of a negative slope is unlikely near the point where a zero with positive slope is expected.

If $\langle G(x)\rangle_{Av}=\bar\xi$ has a null at $x=x_0$, the function is conveniently represented by means of a Taylor expansion about that point.

$$\bar\xi=G(x_0)+\left[\frac{\partial}{\partial x}\langle G(x)\rangle_{Av}\right]_{x=x_0}(x-x_0)=\bar\eta_0(x-x_0) \tag{49}$$

where

$$\bar\eta_0\equiv\left[\frac{\partial}{\partial x}\langle G(x)\rangle_{Av}\right]_{x=x_0}.$$

Similarly,

$$\bar\eta\cong\bar\eta_0+\left.\frac{\partial\bar\eta}{\partial x}\right|_{x=x_0}(x-x_0). \tag{50}$$

Substituting Eqs. (49) and (50) into Eq. (48), one obtains

$$p(x)\cong\frac{1}{(2\pi\mu_{11})^{\frac{1}{2}}}\left\{\bar\eta_0+\left[\left.\frac{\partial\bar\eta}{\partial x}\right|_{x=x_0}-\frac{\mu_{12}}{\mu_{11}}\bar\eta_0\right](x-x_0)\right\} \exp\left[-\frac{(x-x_0)^2}{2\frac{\mu_{11}}{\bar\eta_0^2}}\right]. \tag{51}$$

If the system is to be accurate, $(\mu_{11}/\bar\eta_0^2)^{\frac{1}{2}}$ must be very small compared to x_0. Then the exponential term differs from zero appreciably only for values of x close to x_0, and one therefore feels that the term

$$\left[\left.\frac{\partial\bar\eta}{\partial x}\right|_{x=x_0}-\frac{\mu_{12}}{\mu_{11}}\bar\eta_0\right](x-x_0)$$

will be negligible compared to $\bar\eta_0$. Numerical computation establishes easily that this is indeed the case for the order of magnitude of the parameters under discussion. Under the same conditions μ_{11} does not differ significantly from its value μ_{11_0} at x_0. The distribution is then essentially normal and has the variance

$$\frac{\mu_{11_0}}{\bar\eta_0^2}=\frac{\mu_{11_0}}{S^2}.$$

[10] If the spectrum becomes extremely narrow, the problem approximates that of the automatic frequency control circuit, solutions for which are well known.

[11] S. O. Rice, Bell System Tech. J. **23**, 282 (1944) and **24**, 46 (1945).

[12] See reference 11, p. 51.

JOURNAL OF APPLIED PHYSICS VOLUME 25, NUMBER 8 AUGUST, 1954

On the Distributions of Signals and Noise after Rectification and Filtering

M. A. MEYER, *Laboratory for Electronics, Inc., Boston 14, Massachusetts*

AND

DAVID MIDDLETON, *Cruft Laboratory, Harvard University, Cambridge 38, Massachusetts*

(Received January 9, 1954)

The probability distributions for broad- and narrow-band signals and normal random noise, following a square-law rectifier *and* a video (or audio) filter of arbitrary width, are examined. The present approach is based on a method originally used by Kac and Seigert [J. Appl. Phys. 18, 383 (1949)] for noise alone (after quadratic rectification but no filter) in the case of the nth-order distribution ($n \geq 2$). The procedure requires an appropriate transformation to express the output waveform in terms of the input disturbance; the statistics of the output are now determined by a suitable transformation with respect to the original, normal statistics. Explicit solutions are obtained for an integral equation involving the autocorrelation function of the noise, the weighting function of the video (or audio) filter, and an orthonormal set of eigenfunctions. For noise alone only the eigenvalues need be determined, but for a signal and noise it is necessary to know the eigenfunctions as well. Among the new results are (1), the calculation of the general, nth-order characteristic function for the filtered output following quadratic rectification for an input noise and a signal *and* noise; (2), a discussion of the second-order probability density W_2 in this case; (3), some examples of practical interest; and (4), the limiting cases of broad and narrow post-detection filters (vis-à-vis the predetection filter), with particular attention to approximations to W_2 for signal and noise for narrow videos and an improved approximation for W_1 for noise alone, as well.

1. INTRODUCTION

IN most electronic systems we are concerned with the observation of some physical quantity. This quantity may be random, or it may have a well-defined representation in time and may be accompanied by noise, which can interfere with the measurement. Here we consider as typical inputs to our electronic system either normal random noise, or a normal random signal (the quantity of interest), or a combination of normal random noise and a signal with a definite structure. It is assumed, in addition, that the random processes are stationary, and in some cases, ergodic.

To obtain the desired information from the input, a generalized filter is used which may be active or passive and which may contain both linear and nonlinear filters. One such generalized filter is the combination of an IF amplifier followed by a quadratic detector, which in turn precedes a second, linear ("video") filter. Arrangements of this type are common in radar receivers, for example, and are also basic for power-spectrum analyzers. In the present paper we attempt to determine the statistical properties of the *video filter's* output, since from these statistics the necessary description of the observations can be given, and many of the observed properties of the original wave obtained.[1-3]

A random process can be completely described by a set $W_1, \cdots, W_n$ of probability (density) functions.[4] Here W_n is the nth-order probability density function $W_n(X_1; X_2,t_2; \cdots; X_n,t_n)$ which gives the joint probability of finding an n-tuple of values of X in the ranges $(X_1,X_1+dX_1), \cdots, (X_n,X_n+dX_n)$, at the successive times $t_1(=0), t_2, t_3, \cdots$, etc. Since the process is assumed stationary, only the time differences between pairs of values of X are significant. An important consequence of the normal structure of the input wave is that the set $W_1, \cdots, W_n$ is completely specified if the (co-) variances of the input wave are known. (Here the means are assumed equal to zero.) Thus, if the statistics of the input wave are given, it should be possible in principle to calculate the statistics anywhere else in the system.

The problem considered here is the effect on the statistics of the normal process, when the incoming random wave is subjected to a linear filtering operation, represented by the IF amplifier (Fig. 1), rectified by a quadratic detector, and then passed again through a second linear filter, which we shall call the "video." In particular, we wish the second-order probability density W_2 for the output of this second linear filter—an output which will not in general possess normal properties, because of the intervening nonlinear operation. We use here a generalization of a method first introduced by Kac and Siegert,[5] who obtained the first-order density W_1 for noise and for a signal and noise. Before this, earlier work in the unfiltered (i.e., infinitely wide-video) case[1,2] has given us W_1 and W_2 in a direct way, for all nonlinear elements. Kac and Siegert showed, however, for the first time, that explicit results in the much more difficult situation of arbitrary post-detection filtering could be obtained, at least for the full-wave quadratic rectifier. Their results were subsequently verified by Jastram[6] in certain instances. It turns out

[1] S. O. Rice, Bell System Tech. J. 23, 282 (1944); 24, 46 (1945).
[2] D. Middleton, Quart. Appl. Math. 5, 445 (1948); see also J. Appl. Phys. 22, 1143, 1153 (1951).
[3] Davenport, Johnson, and Middleton, J. Appl. Phys. 23, 377 (1952).
[4] M. C. Wang and G. E. Uhlenbeck, Revs. Modern Phys. 17, 323 (1946).
[5] M. Kac and A. J. F. Siegert, J. Appl. Phys. 18, 383 (1949).
[6] P. Jastram, "The effect of nonlinearity and frequency distortion on the amplitude distribution for stationary random processes," doctoral dissertation, University of Michigan (1947).

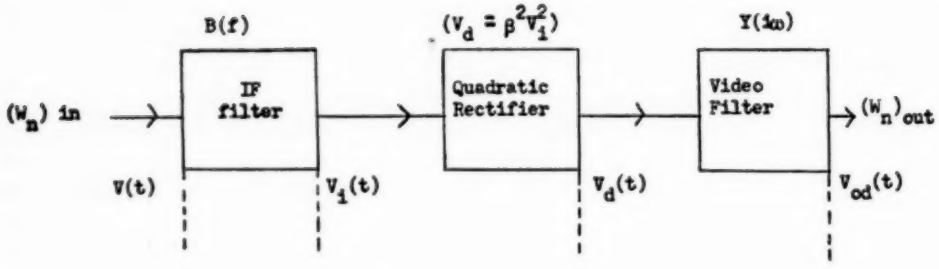


FIG. 1. The generalized filter.

that W_1 can be represented explicitly,[7] provided the eigenvalues of a certain integral equation can be found. For signals and noise both the eigenvalues and the associated eigenfunctions of this integral equation are needed, and, moreover, only the characteristic function of the distribution is expressible in convenient form. Approximations for W_1 have also been developed in the two extremes of video filters wide or narrow compared with the first, or IF response; see Fig. 1.

The choice of a quadratic rectifier is important for several reasons: (1) In receiving systems which must reproduce adequately over extended dynamic ranges, the minimum (or threshold) signal, and therefore the noise, is confined to the lower portions of the rectifier characteristic, which for a wide class of practical detectors approximates a square-law[8] response. (2) Often one desires a quantity proportional to the instantaneous power in a random wave, and this requires usually a quadratic rectifier. (3) From a purely mathematical point of view, the quadratic rectifier is much easier to treat analytically than other nonlinear devices, and in fact, seems at the moment to be the only response for which more or less explicit solutions can be found. (4) Finally, a study of W_2 (as well as W_1) after rectification and filtering provides results useful in the problem of determining, say, the correlation function or spectrum of the output of a *second* nonlinear device, following the video filter. Methods to date have assumed that any filters between successive nonlinear operations are either very broad or very narrow, in which case standard procedures[1,2] may be used.

Kac and Siegert achieved their results in two ways. Their first approach, which we shall call the "expansion method," develops the input random wave in a series using an appropriate set of orthonormal functions. However, this method does not appear possible in higher-order problems (W_2, etc.), since the description by such a set is not general enough.[9] The second, or "direct," derivation employed by Kac and Siegert is the basis of our approach to the higher-order problems. This direct approach requires an appropriate transformation to express the output wave form in terms of the input. The statistics of the output are then determined by suitable additional transformations with respect to the original, input statistics. With this method Kac and Siegert also obtained the nth-order characteristic function for the detector's output in the case of noise alone.

The present paper obtains the following new results:

(1) The general nth-order characteristic function for the filtered output following quadratic rectification, in terms of the eigenvalues (and eigenfunctions) of a certain integral equation, *for signal and noise*, as well as noise alone.

(2) The second-order characteristic function $F_v(\zeta_1,\zeta_2)$ as a special case of (1) for both broad- and narrow-band input waves.

(3) A special video filter which gives the same second-order density W_2 as for an infinitely wide video, in the case of a band-limited IF response.

(4) Approximations for W_2 in the limiting cases of wide-or narrow-videos for noise alone, and for a signal and noise.

These results show, as expected, that the distribution tends to that obtained for an infinitely wide filter when the final filter is wide compared to the IF, and to a normal distribution,[5,9a] when the final filter is narrower than the IF.

The main features of the analysis, which is necessarily quite involved, are outlined in the succeeding sections, along with a summary of special results of interest in the limiting cases. Sufficient detail is maintained to facilitate calculations of any additional cases, if so desired.

2. THE CHARACTERISTIC FUNCTION AFTER RECTIFICATION. NOISE ALONE

Our first major problem is to find the nth-order characteristic function ϕ_n for the output of the quadratic rectifier in terms of the statistics of the input IF wave. We next determine the transformed characteristic function for the output of a linear filter when the input has a given characteristic function. Applying these operations successively gives us the desired charac-

[7] Recent work of R. C. Emerson, J. Appl. Phys. **24**, 1168 (1953), discusses this problem from an alternative point of view, using the cumulants (semi-invariants) of the distribution to avoid solving the associated integral equations of Kac and Siegert's and the present approach.

[8] J. H. Van Vleck and D. Middleton, J. Appl. Phys. **17**, 940 (1946); see also D. Middleton, Proc. Inst. Radio Engrs. **36**, 1467 (1948).

[9] M. A. Meyer, "On Some Distribution Functions in the Theory of Random Noise," doctoral dissertation, Harvard University, June (1952), pp. 2–6, 2–7, and following.

[9a] G. R. Arthur, J. Appl. Phys. **23**, 1143 (1952).

teristic function $F_V(\zeta_1,\cdots,\zeta_n)$ for the output of the video filter when the input is the detected IF disturbance.

Two forms of F_V are needed in limiting cases. One of these is expressed in terms of the eigenvalues and eigenfunctions of the integral equation mentioned above; the other is associated with the Fredholm determinant of this integral equation. The quantity F_V is calculated initially in terms of the time-response (i.e., weighting) functions of the IF and video filters. We next transform these expressions of F_V into representations where the frequency-response (i.e., system) functions now appear. From the integral equation and the determinantal expansion in this new form we obtain finally the desired approximations and other results, since a "wide" or "narrow" video has meaning primarily in the frequency domain.

We start with the nth-order normal distribution density and associated characteristic function of a random variable x which has a zero mean:[10]

$$W_n(x_1,\cdots,x_n)=(2\pi)^{-n}|\mathfrak{y}|^{-1/2}\times\exp\left\{-(1/2)\sum_{kl}^{n}\frac{\mu^{kl}}{|\mathfrak{y}|}x_kx_l\right\}, \quad (2.1a)$$

$$\phi(\zeta_1,\cdots,\zeta_n)\equiv\langle\exp(i\sum_k\zeta_kx_k)\rangle=\exp(-(1/2)\sum_{kl}\zeta_k\zeta_l\mu_{kl}), \quad (2.1b)$$

where $\mathfrak{y}$ is the matrix of the variances $\langle x_kx_l\rangle$, $\langle x_k,x_l\rangle=0$; μ^{kl} is the cofactor of the kth row, lth column of $\mathfrak{y}$, and $\mathfrak{y}$ itself is a symmetric matrix, since $\langle x_kx_l\rangle=\langle x_lx_k\rangle$. Furthermore, we assume an ergodic process, so that the statistical averages $\langle x_kx_l\rangle$ are equivalent to the time averages $\langle x_kx_l\rangle_{Av}=R(t_k-t_l)$, $(k,l=1,\cdots,n)$, where the R_{kl} form a set of autocorrelation functions, and $x_1\cdots x_n$ represent the random wave at n successive times $(t_1,\cdots,t_n)$. By the theorem of Wiener and Khintchine[11,12] we can write

$$R(\tau)=\int_0^\infty w(f)\cos2\pi\tau df, \quad (2.2)$$

where $w(f)$ is the (power) spectral density of the random wave (in this case before it enters the IF filter). Thus, if we know the spectral density, we can completely specify the probability density (2.1) in this instance.

When the input is passed through the IF, one easily shows that the correlation function and spectrum of the output are

$$\langle V_i(t+t_j)V_i(t+t_k)\rangle=R(t_k-t_j)_{IF}=\int_0^\infty w(f)_{IF}\cos2\pi f(t_k-t_j)df, \quad (2.2a)$$

and

$$w(f)_{IF}=|B(f)|^2w(f), \quad (2.2b)$$

where $B(f)$ is the system function of the IF filter. In particular, if the input wave is white noise of constant spectral intensity D, then $w(f)_{IF}=D|B(f)|^2$; henceforth $V_i(t)$ represents the random input to the quadratic detector.

We now distinguish between various outputs of the IF amplifier on the basis of their spectral character: when the power spectrum (measured, say, between half-intensity points) is narrow compared to the central or "resonant" frequency of the wave, it is called *narrow-band*; when this condition is not satisfied, we have *broad-band* noise. The former, for instance, is typical of white noise through a relatively high-Q filter. It turns out analytically that it is more convenient to discuss first these narrow-band cases, and then to specialize the results to the broad-band situation. Accordingly, we begin by assuming a narrow-band wave possessing a symmetrical spectrum, centered at the frequency f_0, as shown in Fig. 2. The wave $V_i(t)$ can then be represented as a combination of slowly and rapidly varying components in the usual way[1,2]

$$V_i(t)=R(t)\cos[\omega_0t+\Psi(t)]=x(t)\cos\omega_0t+y(t)\sin\omega_0t, \quad (2.3a)$$

$$x(t)=R\cos\Psi;\quad y(t)=R\sin\Psi, \quad (2.3b)$$

where R and Ψ are, respectively, the envelope and phase of the wave.

The quadratic detector through which $V_i(t)$ is next passed yields the instantaneous output

$$V_0(t)=\beta^2V_i(t)^2,\quad(-\infty<V_i<\infty), \quad (2.4a)$$

$$=\frac{\beta^2}{2}[(x^2+y^2)+(1/2)(x^2\cos2\omega_0t+(2xy-y^2)\sin2\omega_0t)] \quad (2.4b)$$

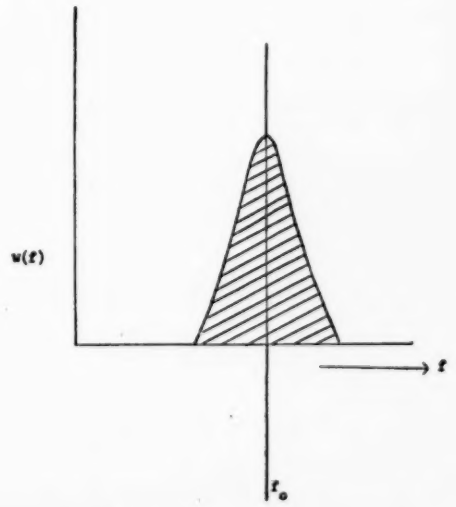


FIG. 2. Narrow-band noise spectrum.

[10] H. Cramér, *Mathematical Methods of Statistics* (Princeton University Press, Princeton, (1946); see Chap. 10.
[11] N. Wiener, Acta Math. **55**, 117 (1930).
[12] A. Khintchine, Math. Ann. **109**, 604 (1934).

for narrow-band waves. (For simplicity, in subsequent work we set $\beta^2=2$ for narrow-band waves and $\beta^2=1$ for broad-band disturbances, the extension to other values being obvious.) The low-frequency output, with which we are concerned here, is at once

$$V_0(t)_{LF}=x^2+y^2;\ \text{at}\ t=t_j,$$
$$(V_{LF})_j=x(t_j)^2+y(t_j)^2=x_j^2+y_j^2, \tag{2.5}$$

this last for the n time points t_j $(j=1\cdots n)$. But since $V_i(t)$ is normally distributed, so also are x and y,[1,2] giving the moments

$$\langle x_j^2\rangle=\int_0^\infty w(f)_{IF}df=\langle y_j^2\rangle\doteq\int_0^\infty w(f-f_0)_{IF}df\equiv\psi/2,$$
$$=\langle V_N^2\rangle, \tag{2.6a}$$

$$\langle x_jx_k\rangle=\langle y_jy_k\rangle\doteq\int_0^\infty w(f-f_0)_{IF}\cos\omega(t_k-t_j)df$$
$$\equiv\frac{\psi}{2}\alpha(t_k-t_j)=\frac{\psi}{2}\alpha_{kj}, \tag{2.6b}$$

$$\langle x_jy_k\rangle=-\langle x_ky_j\rangle\doteq\int_0^\infty w(f-f_0)_{IF}\sin\omega(t_k-t_j)df, \tag{2.6c}$$

this last since the spectrum is narrow-band.

The nth-order characteristic function associated with the nth-order distribution density *of the low-frequency output* of the quadratic detector at n successive times is

$$\phi_N(\zeta_1,\cdots,\zeta_n)\equiv\left\langle\exp i\sum_{j=1}^n\zeta_jV_{LF}(t+t_j)\right\rangle$$
$$=\int_{-\infty}^\infty\cdots\int W_n(x_1\cdots x_n;y_1\cdots y_n)$$
$$\times\left[\exp\left\{i\sum_{j=1}^n\zeta_j(x_j^2+y_j^2)\right\}\right]dx_1\cdots dx_ndy_1\cdots dy_n, \tag{2.7}$$

where W_n is the $2n$th-order distribution obtained from (2.1) on letting

$$x_k=x_k,\ (k,l=1,\cdots,n);$$
$$x_k=y_k,\ (k,l=n+1,\cdots,2n)$$

therein. In most instances we can assume a symmetrical pass band, so that $\langle x_jy_k\rangle$, Eq. (2.6c), vanishes, whereupon $W_n(x_1,\cdots,x_n;y_1,\cdots,y_n)$ factors into $W_n(x_1\cdots x_n)\times W_n(y_1\cdots y_n)$, each term of which is precisely (2.1). With the notation of Kac and Siegert,[5] *viz.*, (2.6b):

$$\langle x_jx_k\rangle=(1/2)\int_{-\infty}^\infty w(f)_{IF}\cos(\omega-\omega_0)(t_k-t_j)df=\frac{\psi}{2}\alpha_{jk},$$

we obtain finally for independent x and y

$$\phi_N(\zeta_1,\cdots,\zeta_n)=(\psi^n\pi^n|\alpha|)^{-1}\left[\int_{-\infty}^\infty\cdots\int dx_1\cdots dx_n\right.$$
$$\left.\times\exp\left\{-\sum_{k,l=1}^n\left(\frac{\alpha^{kl}}{|\psi\alpha|}-i\zeta_k\delta_{kl}\right)x_kx_l\right\}\right] \tag{2.8}$$

in which δ_{kl} is the familiar Kronecker delta: $\delta_{kl}=1$, $k=l$; $=0$, $k\neq l$.

The next step is to integrate (2.8) by reducing the quadratic form to the sum of squares (see Appendix I). When this is done we obtain a result of Kac and Siegert[5]:

$$\phi_N(\zeta_1,\cdots,\zeta_n)=i^n/|\zeta_k\psi\rho_{kj}+i\delta_{kj}|,\quad \rho_{kj}=\alpha_{kj}, \tag{2.9}$$

where we have replaced α_{kj} by ρ_{kj} to get a more familiar nomenclature. The characteristic function ϕ_N is given in terms of the determinant $|\zeta_k\psi\rho_{kj}+i\delta_{kj}|$, in which the statistical nature of the input noise enters only through the correlation coefficients

$$\psi\rho_{kj}=\int_{-\infty}^\infty w(f-f_0)_{IF}\cos\omega(t_k-t_j)df=2\langle x_jx_k\rangle.$$

An alternative form of (2.9), which is more useful in the discussion to follow, is (see Appendix I):

$$(\textit{narrow-band}):\begin{cases}\phi_N(\zeta_1,\cdots,\zeta_n)=\prod_{j=1}^n(1-i\lambda_j)^{-1}, & \text{(2.10a)}\\ \text{with} & \\ \psi\sum_{k=1}^n\zeta_k\rho_{km}f_{jk}=\lambda_jf_{jm}, & \text{(2.10b)}\end{cases}$$

i.e. λ_j is the jth eigenvalue of the set of homogeneous linear equations given in (2.10b), and f_{jk} are the components of the jth eigenvector f_j. Note that in this representation λ_j is a function of $\zeta_1,\zeta_2,\cdots,\zeta_n$.

The broad-band case for noise alone follows at once from the observation that now the output of the detector (2.4a) can be written $V_0(t)=\beta^2X^2(t)$, where the moments are

$$\langle X_jX_k\rangle=\psi\rho_{jk}=\int_{-\infty}^\infty w(f)_{IF}\cos\omega(t_j-t_k)df;$$
$$\langle X_j^2\rangle=\langle X_k^2\rangle=\psi,\ \text{etc.} \tag{2.11}$$

Setting $\beta=1$ here, and repeating the procedure outlined above for (2.7)–(2.9), we obtain just the square root of (2.9) and (2.10) for the characteristic function after rectification in this case:

$$(\textit{broad-band}):\phi_N(\zeta_1,\cdots,\zeta_N)=i^{n/2}/|\zeta_k\psi\rho_{kj}+i\delta_{kj}|^{1/2}$$
$$=\prod_{j=1}^n(1-i\lambda_j)^{-1/2}. \tag{2.12}$$

3. CHARACTERISTIC FUNCTION AFTER RECTIFICATION. SIGNAL AND NOISE

As in Sec. 2, we divide the discussion into two parts, considering first the more general analysis of the narrow-band case, where now a signal as well as noise constitutes the input to the IF filter. The noise part of the rectifier's input is given as before by (2.3), while the signal is represented by $M(t')\cos\omega_0t$. Here $M(t')$ is a

modulation of the carrier† $\cos\omega_0 t$, and t' rather than t is used to indicate the fact that modulation and (IF) carrier are uncorrelated. There is, of course, no correlation between signal and noise either. The total input to the detector is therefore

$$V_i(t)=[x(t)+M(t')]\cos\omega_0 t+y(t)\sin\omega_0 t, \quad (3.1)$$

and from (2.5) the corresponding low-frequency output of the quadratic detector becomes ($\beta=\sqrt{2}$)

$$V_{LF}(t)=x^2+y^2+2xM+M^2. \quad (3.2)$$

Letting $x(t+t_j)=x_j$, $y(t+t_j)=y_j$, $M(t'+t_j)=M_j(t')$, we find that the characteristic function takes the more general form

$$\phi_{S+N}(\zeta_1,\cdots,\zeta_n)=\int_{-\infty}^{\infty}\cdots\int W_n(x_1,\cdots,x_n)W_n(y_1,\cdots,y_n)$$
$$\times\exp\Big(i\sum_{j=1}^{n}\zeta_j[x_j^2+y_j^2+2x_jM_j(t')+M_j(t')^2]\Big)dx_1\cdots dy_n, \quad (3.3)$$

which is explicitly, see (2.8),

$$\phi_{S+N}(\zeta_1,\cdots,\zeta_n)=(\psi^n\pi^n|\alpha|)^{-1}\exp\sum_{j}^{n}i\zeta_jM_j^2(t')$$
$$\cdot\int_{-\infty}^{\infty}\cdots\int dy_1\cdots dy_n\exp\Big(i\sum_{k}^{n}\zeta_ky_k^2-\sum_{kl}^{n}\frac{\alpha^{kl}y_ky_l}{|\alpha|\psi}\Big)$$
$$\cdot\int_{-\infty}^{\infty}\cdots\int dx_1\cdots dx_n\exp\Big(i\sum_{kl}^{n}\zeta_kx_k^2-\sum_{kl}^{n}\frac{\alpha^{kl}x_kx_l}{\psi|\alpha|}+2i\sum_{k}^{n}\zeta_kx_kM_k(t')\Big). \quad (3.4)$$

Again introducing a suitable principal-axis transformation (see Appendix I) and integrating give us finally

$$(\textit{narrow-band}):\ \phi_{S+N}(\zeta_1\cdots\zeta_n)=\prod_{k=1}^{n}(1-i\lambda_k)^{-1}$$
$$\cdot\prod_{k=1}^{n}\exp\Big[i\Big(\sum_{j}^{n}M_j(t')\zeta_jf_{kj}\Big)^2\Big/(1-i\lambda_k)\Big], \quad (3.5a)$$

where the eigenvalues $\lambda_j=\lambda_j(\zeta_1,\cdots,\zeta_n)$ and eigenfunctions f_{kj} are defined by (2.10b), with the additional normalizing relation

$$\sum_{l=1}^{n}\zeta_lf_{kl}f_{jl}=\delta_{kl}. \quad (3.5b)$$

The corresponding expression for broad-band waves is easily found. We now let $S(t')$ represent the signal and $X(t)$ the noise, so that the output (2.4a) of the quadratic detector, when $V_i=X+S$ is the input, becomes at the times t_j ($j=1,\cdots,n$)

$$V_j=[X_j+S_j(t')]^2=X_j^2+S_j^2(t')+2X_jS_j(t'), \quad (\beta=1). \quad (3.6)$$

Comparison with (3.2) and (3.4) gives at once the desired characteristic function

$$(\textit{broad-band}):\ \phi_{S+N}(\zeta_1,\cdots,\zeta_n)=\prod_{k=1}^{n}(1-i\lambda_k)^{-1/2}$$
$$\cdot\prod_{k=1}^{n}\exp\Big[i\Big(\sum_{j=1}^{n}S_j(t')\zeta_jf_{kj}\Big)^2\Big/(1-i\lambda_k)\Big]. \quad (3.7)$$

Both (3.5a) and (3.7) are generalizations of Kac and Siegert's earlier result (2.12) for noise alone.

4. THE POST-DETECTION FILTER

There remains now to determine the effects of the linear, post-detection ("video") filter on the statistics of the rectified wave. Our first task is to obtain the desired characteristic function after this post-detection filter for the broad and narrow-band cases discussed above. Special limiting cases of interest are reserved for Secs. (5–7) and Appendix II.

If the input to the "video" filter is $V_d(t)$ where $V_d(t)$ is $V_{LF}(t)$, for broad- or narrow-band inputs, and if $V_{0d}(t)$ is the corresponding output of the filter, we can write in the steady state

$$V_{0d}(t)=\int_{-\infty}^{\infty}A(\tau)V_d(t-\tau)d\tau, \quad (4.1)$$

where $A(\tau)$ is the filter's weighting function, which is related to the system function $Y(i\omega)$ by the Fourier transforms

$$A(t)=\int_{-\infty}^{\infty}Y(i\omega)e^{i\omega t}df, \quad Y(i\omega)=\int_{-\infty}^{\infty}A(t)e^{-i\omega t}dt \quad (4.2)$$

for linear stable filters. For all real filters $A(t)=0$, $t<0-$.

The procedure now is to replace the integral (4.1) by a sum which (in the appropriate limit) will again be equal to the integral. The sum is here

$$(V_{0d})_k\equiv V_k=\sum_{l=0}^{L}A_l(V_d)_{(k-l)},$$

and

$$(V_{0d})_{k-m(j)}=\sum_{l=0}^{L}A_l(V_d)_{k-l-m(j)}, \quad (4.3)$$

where the interval over which $A(\tau_l)$ is significant is divided into $L+1$ points, each of which is a distance h apart, and for each $(k-1)h\leq\tau_k\leq kh$; $A(\tau_l)=A_lh$.

† Since we are dealing with stationary waves, one can always choose the origin of time such that $\cos(\omega_0t'+\phi)=\cos\omega_0t$.

The nth-order characteristic functions $F_V(\zeta_1', \cdots, \zeta_n')$ for the output $V_{0d}(t)$ of the post-detection filter is

$$F_V(\zeta') = \left\langle \exp\left\{ i \sum_{k=1}^{n} \zeta_k V_{0d}(t-\tau_{k-1}) \right\} \right\rangle. \quad (4.4)$$

Substituting for $V_{0d}(t-\tau_j)$, etc., from (4.3) into (4.4) and performing the statistical average gives

$$F_V(\zeta') = \int_{-\infty}^{\infty} \cdots \int W_L[(V_d)_0, \cdots, (V_d)_{(-L)}] \times \exp\left\{ i \sum_{k=1}^{n} \zeta_k \sum_{l=0}^{L} A_l (V_d)_{(k-l-m(k-1))} \right\} \times d(V_d)_0 \cdots d(V_d)_{-L}, \quad (4.5)$$

where $[m(j)-1]h \leq \tau_j \leq m(j)h$. Here W_L is the $(L+1)$st-order probability density for the $L+1$ inputs to the filter at the successive times $t-\tau_L$ to t. (According to the present notation $m(0)=0$.)

Our next step is to replace‡ the sum

$$\sum_{l=0}^{L} A_l (V_d)_{(k-l-m(j))} \quad \text{by} \quad \sum_{l=0}^{L} A_{l-m(j)} (V_d)_{k-l}.$$

Equation (4.5) is now written with the help of the Fourier transform of W_n, *viz.*, the characteristic function $\phi(\zeta_0, \cdots, \zeta_{-L})$,

$$F_V(\zeta') = \int_{-\infty}^{\infty} \cdots \int d(V_d)_0 \cdots d(V_d)_{-L} \times \exp\left\{ i \sum_{k=1}^{n} \zeta_k' \sum_{l=0}^{L} A_{l-m(k-1)} (V_d)_{k-l} \right\} \cdot (2\pi)^{-L} \int_{-\infty}^{\infty} \cdots \int d\zeta_0 \cdots d\zeta_{-L} \phi(\zeta_0, \cdots, \zeta_{-L}) \times \exp\left\{ -i \sum_{l}^{n} \zeta_{-l} (V_d)_{-l} \right\}. \quad (4.6)$$

The characteristic function $\phi(\zeta_0, \cdots, \zeta_{-L})$ is associated with the $L+1$ inputs $(V_d)_0, \cdots, (V_d)_{-L}$. Integrating over $(V_d)_{-q}$ $(q=0, \cdots, L)$ we find the first set of integrals to be

$$\sum_{l=0}^{L} \delta\left(\zeta_{-l} - \sum_{k=1}^{n} \zeta_k' A_{l-m(k-1)} \right) \quad (4.6a)$$

so that (4.6) becomes finally

$$F_V(\zeta') = \phi\left(\sum_{k=1}^{n} \zeta_k' A_{-m(k-1)}; \cdots; \sum_{k=1}^{n} \zeta_k' A_{j-m(k-1)}; \cdots; \sum_{k=1}^{L} \zeta_k' A_{L-m(k-1)} \right), \quad (4.7)$$

where we remember that ϕ is here the characteristic function for the *output* of the quadratic rectifier, and $F_V(\zeta')$ is the desired characteristic function of the output of the rectifier plus post-detection filter. A number of specific cases follow:

(a) *Narrow-Band Noise*

We apply (4.7) now, using the results of Secs. 2 and 3. Equations (2.9) and (4.7) give us first one form of the nth-order characteristic function for narrow-band noise after video filtering:

$$F_V(\zeta')_N = i^{L+1} \Big/ \left| \psi \rho_{kj} \left(\sum_{q=1}^{n} \zeta_q' A_{k-m(q)} \right) + i\delta_{kj} \right|, \quad (k=1, \cdots, L), \quad (4.8)$$

and the determinant is an $(L \times L)$ one. Corresponding to (2.10) we have the alternative representation

$$F_V(\zeta)_N = \prod_{j=0}^{L} (1-i\lambda_j)^{-1}, \quad (4.9a)$$

where

$$\sum_{k=0}^{L} \psi \rho_{km} \left(\sum_{q=1}^{n} \zeta_q' A_{k-m(q-1)} \right) f_{jk} = \lambda_j f_{jm}, \quad (4.9b)$$

and since $A_l = A(l_l)h$ this sum can be rewritten

$$\left\{ \sum_{k=0}^{L} (\psi \rho_{km} f_{jk}) \sum_{q=1}^{n} \zeta_q' A(l_k - l_{m(q-1)}) \right\} h = \lambda_j f_{jm}. \quad (4.9c)$$

Letting L approach ∞ and h approach 0 we get an integral for the sum over k, and so for the characteristic function of the continuous, filtered wave $V_{0d}(t)$, Eq. (4.1), the result (4.9a) becomes in the limit (dropping the primes on ζ_1, etc.)

$$F_V(\zeta)_N = \prod_{j=1}^{\infty} (1-i\lambda_j)^{-1}, \quad (4.10a)$$

where λ_j are the eigenvalues (here functions of $\zeta_1, \cdots, \zeta_n$ and $\tau_1, \cdots, \tau_{n-1}$) of the integral equation§

$$\int_0^{\infty} dt \left[\sum_{q=1}^{n} \zeta_q A(t-\tau_{q-1}) \right] \rho(t,s) f(t) = \lambda f(s). \quad (4.10b)$$

‡ This substitution may be justified by the following argument: if we let $l+m(j)=q$, then

$$\sum_{l=0}^{L} A_l (V_d)_{k-l-m(j)} = \sum_{m(j)}^{L+m(j)} A_{(q-m(j))} (V_d)_{k-q} = \sum_{0}^{L+m(j)} A_{(q-m(j))} (V_d)_{k-q},$$

since $A(\tau_j)=0$; $\tau_j<0$. Thus, if L is chosen sufficiently large, so that $A_{(L-m(j))} \to 0$, then $A_L \to 0$ also, and we can replace the upper limit on the last sum with L only.

§ The eigenvalues of the finite sum are certainly discrete, but it may be questioned whether or not they remain discrete in the limit of the integral equation, since one can always find kernels for which this is not the case. However, for symmetric kernels

In the limit, of course, one gets an infinite determinant, if (4.8) is used.

The examples of chief interest occur for $n=1, 2$. For $n=1$ (first-order probability density and characteristic function) we find that (4.10b) reduces to

$$\psi\zeta_1\int_0^\infty dtA(t)\rho(t,s)f(t)=\lambda f(s)=\lambda(\zeta_1)f(s). \quad (4.11a)$$

Letting $\lambda_j=\lambda_j'\zeta_1\psi$, we see that the integral equation is modified to

$$\int_0^\infty A(t)\rho(t,s)f(t)dt=\lambda' f(s), \quad (4.11b)$$

and the desired characteristic function becomes

$$F_V(\zeta_1)_N=\prod_{j=1}^{\infty}(1-i\zeta\psi\lambda_j')^{-1}, \quad (4.12)$$

which is the result obtained by Kac and Siegert[5] when $\psi=1$.

When $n=2$, we get from (4.10a, b)

$$F_V(\zeta_1,\zeta_2;\tau)_N=\prod_{j=1}^{\infty}[1-i\lambda_j(\zeta_1,\zeta_2;\tau)]^{-1} \quad (4.13a)$$

with the accompanying integral equation

$$\psi\int_0^\infty dt[\zeta_1A(t)+\zeta_2A(t-\tau)]\rho(t,s)f(t)=\lambda f(s). \quad (4.13b)$$

The alternative form (4.8) proves useful in limiting cases of wide and narrow video. If we let

$$\Delta=|\delta_{jk}+\gamma hK_{jk}|$$

be the Lth-order determinant (4.8), where $A_k=A(t_k)h$, we may expand Δ as follows[13]

$$\Delta=1-\gamma\sum_j^L K_{jj}h+\frac{\gamma^2}{2}\sum_{jk}^L h^2\begin{vmatrix}K_{jj} & K_{jk}\\ K_{kj} & K_{kk}\end{vmatrix}\cdots. \quad (4.14)$$

In the limits $M\to\infty$, $h\to 0$, sums are replaced by integrals, as before, $\gamma=i$, and we get finally

$$F_V(\zeta_1,\zeta_2;\tau)_N=\left\{1-i\int_0^\infty K(t,t)_\tau dt-(1/2!)\right.$$

$$\left.\times\int_0^\infty\int\begin{vmatrix}K(t_1,t_1)_\tau & K(t_1,t_2)_\tau\\ K(t_2,t_1)_\tau & K(t_2,t_2)_\tau\end{vmatrix}dt_1dt_2+\cdots\right\}^{-1}, \quad (4.15a)$$

where

$$K(t_1,t_2)_\tau=\psi\rho(t_2-t_1)[\zeta_1A(t_1)+\zeta_2A(t_1-\tau)]. \quad (4.15b)$$

Applications of this approach are considered in Secs. 5 and 6.

We note finally that if the video filter's weighting function is such that $A(t)\geq 0$, all $t\geq 0$, then only positive eigenvalues can occur, corresponding physically to the fact that the output of such a filter is always positive (when the preceding detector is quadratic). Consequently, there will be a dc term, appearing as a finite average value for the output, i.e., $\langle V_{out}\rangle>0$, as well as an ac fluctuation about this value, but no $\langle V_{out}\rangle<0$. A simple example arises when the video filter is an RC network, with the output taken across the condenser, as then $A(t)=(RC)^{-1}e^{-t/RC}$, $t>0$. However, when $A(t)$ can be negative for some $t(>0)$, a [finite¶] number of negative eigenvalues appear as well: the output can be negative part of the time, although there can still be a positive mean value $\langle V\rangle$ (which vanishes in the limit of a very narrow video). An example of this is provided by a CR post-detection filter, with the output observed now across the resistance, as here

$$A(t)=\delta(t-0)-(RC)^{-1}e^{-t/RC}, \quad (t>0-).$$

These phenomena are specifically illustrated by the results of Secs. 5, 6, see (5.3), (5.4), and (5.22), etc., and the foregoing remarks apply also for (b)–(d) following.

(b) *Broad-Band Noise*

From (2.11), (2.12) it is clear that the results of Sec. 4 may be used directly for broad-band problems, with the

$K(s,t)$, which are quadratically integrable in the sense

$$\int_0^\infty\int K^2(s,t)dsdt=A \text{ (exists)}, \quad \text{(i)}$$

the proof of discreteness is easily given. We summarize the main points of the argument: let the eigenfunctions and eigenvalues be defined from

$$\int_0^\infty K(s,t)y_j(t)dt=\lambda_j y_j(s). \quad \text{(ii)}$$

Since we are looking for quadratically integrable solutions of (ii), we can choose these from an orthonormal set. Now the Fourier coefficients of $K(s,t)$, with respect to $y_j(t)$, are given by $\lambda_j y_j(s)$. Bessel's inequality yields (for arbitrary n)

$$\sum_{j=1}^n \lambda_j^2 y_j^2(s)\leq\int_0^\infty K^2(s,t)dt. \quad \text{(iii)}$$

Integrating over s we obtain

$$\sum_{j=1}^n \lambda_j^2\leq\int_0^\infty\int K^2(s,t)dsdt=A \text{ (exists)} \quad \text{(iv)}$$

from (i). Therefore $\sum_{j=1}^{\infty}\lambda_j^2$ converges, and the λ_j's can have no finite point of accumulation. The spectrum is accordingly discrete. In our examples the quadratic integrability may be shown from the fact that $\int_{-\infty}^{\infty}|A(t)|dt$ exists, and the fact that $\rho(s,t)=\rho(t,s)=\rho(|t-s|)$ never has a weaker falling-off for large $(t-s)$ than $|(t-s)|^P\cdot e^{-|s-t|}$, where P is finite (≥ 0). The former applies because we are dealing with linear, stable filters, and the latter, because ρ represents the autocorrelation function of intensity-limited noise after it has been passed through such filters. We are indebted to Professor Ivar Stakgold for calling to our attention this question of the discrete character of the eigenvalues.

[13] Lovitt, *Linear Integral Equations* (McGraw-Hill Book Company, Inc., New York, 1924), p. 24.

¶ This number depends on the kernel and on the weighting function $A(t)$, of course, but is finite for realizable networks.

simple modifications of Eqs. (4.8), (4.10), (4.13a), etc., i.e. of replacing the right-hand members by their square roots.

(c) *Narrow-Band Signal and Noise*

A direct modification of (3.5a) in conjunction with the transformation used to replace ζ_{-l}' in obtaining (4.7) permits us to write $\zeta_j(\equiv\zeta_{-j})=\zeta_1'A_j$, $(j=0, \cdots, L\to\infty)$ for the first-order characteristic function where there is a signal. Dropping the prime on ζ_1' gives us

$$F_V(\zeta_1)_{S+N}=\prod_{k=1}^{\infty}(1-i\psi\zeta_1\lambda_k')^{-1}$$

$$\times\exp\left\{i\zeta_1^2\left(\int_{-\infty}^{\infty}M(t'+t)A(t)f^{(k)}(t)dt\right)\Big/(1-i\zeta_1\psi\lambda_k')\right\}, \quad (4.16)$$

where $\lambda_k=\lambda_k'\zeta_1\psi$ as in (4.11a, b). The associated integral equation and normalization of the eigenfunctions $f^{(k)}$ follow, respectively, from

$$\int_{-\infty}^{\infty}A(t)\rho(t,s)f^{(k)}(t)dt=\lambda_k'f^{(k)}(s);$$

$$\zeta_1\int_{-\infty}^{\infty}A(t)f^{(k)}(t)^2dt=1. \quad (4.17)$$

From (4.16) it is clear that in dealing with a signal, as well as noise, one must also know the eigenfunctions; it is not enough to determine the eigenvalues alone, as is true of the case for noise only.

The second-order characteristic function is obtained in the same way. The general transformation relating ζ_{-j} and ζ_k' now reduces to the two terms $\zeta_j=\zeta_1'A_j+\zeta_2'A_{j-m}$ $(j=0, \cdots, L\to\infty)$. Dropping primes again we get finally

$$F_V(\zeta_1,\zeta_2;\tau)_{S+N}=\prod_{k=1}^{\infty}(1-i\lambda_k)^{-1}$$

$$\times\exp\left\{i\left(\int_{-\infty}^{\infty}M(t'+t)[\zeta_1A(t)+\zeta_2A(t-\tau)]\cdot f^{(k)}(t)dt\right)^2\Big/(1-i\lambda_k)\right\}, \quad (4.18)$$

where we remember that now $\lambda_k=\lambda_k(\zeta_1,\zeta_2;\tau)$. The associated integral equation and normalization condition on $f^{(k)}$ are

$$\psi\int_{-\infty}^{\infty}[\zeta_1A(t)+\zeta_2A(t-\tau)]\rho(t,s)f^{(k)}(t)dt=\lambda_kf^{(k)}(s), \quad (4.19a)$$

$$\int_{-\infty}^{\infty}[\zeta_1A(t)+\zeta_2A(t-\tau)]f^{(k)}(t)^2dt=1. \quad (4.19b)$$

A check on these results is given in Appendix II in the special case of an infinitely wide video.

(d) *Broad-Band Signal and Noise*

The expressions (4.16)–(4.19) may be adapted at once to this case if we replace $M(t'+t)$ by $S(t'+t)$ and $\Pi_k(1-i\lambda_k)^{-1}$ by $\Pi_k(1-i\lambda_k)^{-1/2}$. [See Eq. (3.7) for infinitely wide video filter; also Appendix II.]

5. THE NARROW VIDEO FILTER

(a) *Noise Alone*

If we consider now the determinantal relation (4.15a) and expand it in such a form as to produce the coefficients of the various powers of $\zeta_1^j\zeta_2^{j-k}$, $\zeta_2^j\zeta_1^{j-k}$, $(j=1\cdots n;\ k=0\cdots j)$ we find that these coefficients consist of a sum of terms of order $(\Delta f_r/\Delta f_F)$, $(\Delta f_r/\Delta f_F)^2$, $\cdots$, $(\Delta f_r/\Delta f_F)^{j-1}$. When the video band width Δf_r is small compared to the IF band width‖ Δf_F a good approximation may be obtained by keeping terms up to and including $(\Delta f_r/\Delta f_F)^2$. Letting $n\to\infty$ and picking out the significant terms to the desired order, we find finally that

$$F_V(\zeta_1,\zeta_2;\tau)_N=e^{i\psi(\zeta_1+\zeta_2)}\left\{1+\frac{\psi^2}{2}[(\zeta_1^2+\zeta_2^2)q(0)+2\zeta_1\zeta_2q(\tau)]+\frac{i\psi^3}{3}[(\zeta_1^3+\zeta_2^3)p(0)+3\zeta_1^2\zeta_2p(\tau)+3\zeta_1\zeta_2^2p(-\tau)]+0(\Delta f_r/\Delta f_F)^3\right\}^{-1}, \quad (5.1)$$

where

$$p(\tau)=\int\int_{-\infty}^{\infty}\int Y(i\omega)Y(i\omega')Y(i\omega'+i\omega)^*\times H(f'')H(f'+f'')H(f-f'')e^{-i\omega\tau}\cdot dfdf'df''\equiv 0(\Delta f_r/\Delta f_F)^2, \quad (5.2a)$$

$$q(\tau)=\int_{-\infty}^{\infty}\int H(f)H(f')|Y(i\omega+i\omega')|^2e^{i(\omega+\omega')\tau}dfdf'=0(\Delta f_r/\Delta f_F). \quad (5.2b)$$

Our next step is to express (5.1) in terms of an exponential with linear and quadratic arguments, the higher-order terms appearing as a factor of the exponential. Inversion then gives the (asymptotic) expression for the second-order probability density for narrow videos (vis-à-vis the IF):

‖ These band widths are those of the equivalent rectangular filters. For details of the method, see reference 8, Chap. V.

(*narrow-band*)

$$W_2(V_1,V_2;\tau)_N\simeq\Big\{1-\frac{\psi^3}{3}\Big[\Big(\frac{\partial^3}{\partial V_1^3}+\frac{\partial^3}{\partial V_2^3}\Big)p(0)+3\frac{\partial^3}{V_1^2\partial V_2}p(\tau)+\frac{3\partial^3}{\partial V_1\partial V_2^2}p(-\tau)\Big]+\frac{\psi^4}{8}\Big[\Big(\frac{\partial^4}{\partial V_1^4}+\frac{\partial^4}{\partial V_2^4}\Big)q(0)+2\frac{\partial^4}{\partial V_1^2\partial V_2^2}(q^2(0)+q^2(\tau))+\Big(\frac{\partial^4}{\partial V_1^3\partial V_2}+\frac{\partial^4}{\partial V_1\partial V_2^3}\Big)4q(0)q(\tau)\Big]+0(\Delta f_\tau^3/\Delta f_F^3)\Big\}W_2(V_1,V_2;\tau)_{N,0},\quad(5.3)$$

where $(W_2)_{N,0}$ is the distribution density defined by

$$W_2(V_1,V_2;\tau)_{N,0}=[2\pi\psi^2q(0)(1-\rho_0^2)^{1/2}]^{-1}\times\exp\Big(-\Big\{\frac{(V_1-\psi)^2+(V_2-\psi)^2-2(V_1-\psi)(V_2-\psi)\rho_0}{2\psi^2q(0)(1-\rho_0^2)}\Big\}\Big).\quad(5.4)$$

in which

$$\rho_0(\tau)\equiv\Phi_1(\tau)/\Phi_1(0)=\int_{-\infty}^{\infty}A(t)A(t+\tau)dt\Big/\int_{-\infty}^{\infty}A^2(t)dt.\quad(5.5)$$

[See Appendix III, Eq. (A3.1).]

For *broad-band noise* one may use (5.3) after inserting a factor $\frac{1}{2}$ before each term in the brackets [ ] and replacing ψ by $\psi/2$, ψ^2 by $\psi^2/2$ in (5.4); ψ itself is then reinterpreted according to the remarks following reference 14.

The desired development of the characteristic function and probability density may also be obtained directly from (4.13a): for the narrow-band case we have formally for small λ_j

$$F_V(\zeta_1,\zeta_2,\tau)_N=\prod_{j=1}^{\infty}[1-i\lambda(\zeta_1,\zeta_2;\tau)]^{-1}=\Delta^{-1}=\exp\Big[-\sum_j^{\infty}\log(1-i\lambda_j)\Big]=\exp\Big(i\sum_j^{\infty}\lambda_j-(1/2)\sum_j^{\infty}(\lambda_j^2+\cdots\Big).\quad(5.6)$$

These sums of the eigenvalues are found in the usual way by expressing the symmetrical kernel (4.15b) as a series in the *orthonormal* set of eigenfunctions $f(t)$

$$K(t_1,t_2)=\textstyle\sum_j\lambda_jf^{(j)}(t_1)f^{(j)}(t_2)$$

and integrating over t_1 and t_2. The sums in (5.6) are easily seen to be the trace of the kernel and of the first iterated kernel, respectively, which in the time and frequency domains become

$$\int_{-\infty}^{\infty}K(t,t),dt=\int_{-\infty}^{\infty}\mathcal{K}(f,f)df=\sum_{j=1}^{\infty}\lambda_j=\psi(\zeta_1+\zeta_2),\quad(5.7)$$

$$\int_{-\infty}^{\infty}\int K(t_1,t_2),K(t_2,t_1),dt_1dt_2=\int_{-\infty}^{\infty}\int\mathcal{K}(f_1,f_2)\mathcal{K}(f_2,f_1)df_1df_2=\sum_{j=1}^{\infty}\lambda_j^2=(\zeta_1^2+\zeta_2^2)\psi^2q(0)+2\zeta_1\zeta_2\psi^2q(\tau),\quad(5.8)$$

and so on for the second and higher iterations. The frequency form of K is

$$\mathcal{K}(f_1,f_2)=\psi(\zeta_1+\zeta_2e^{i(\omega_2-\omega_1)})H(f_1)Y(i\omega_1-i\omega_2).\quad(5.9)$$

The correspondence between these results and (5.1)–(5.2b) indicates that the λ_j are indeed small for the narrow video (5.2), so that this latter approach may be followed for the corresponding problem of broad-band noise and signal and noise (considered below).

(*b*) *Signal and Noise*

Here the governing integral equation and the characteristic function for the second-order case ($n=2$) are given by (4.19), (4.18) for narrow-band waves. As mentioned above, when the video is narrow compared to the IF, the eigenvalues are small and the product $\prod_{k=1}^{\infty}(1-i\lambda_k)^{-1}$ (for noise alone) approaches a normal characteristic function. Using an expansion similar to (5.6) we have for the exponent in (4.18)

$$\exp\{i\sum_1^{\infty}c_k^2-\psi\sum_1^{\infty}\lambda_kc_k^2-i\psi^2\sum_1^{\infty}\lambda_k^2c_k^2+\cdots\},\quad(5.10a)$$

where

$$c_k=\int_{-\infty}^{\infty}M(t'+t)[\zeta_1A(t)+\zeta_2A(t-\tau)]f^{(k)}(\tau)dt.\quad(5.10b)$$

Expanding $M(t'+t)$ in the orthonormal set $f^{(k)}(t)$ gives

$$M(t'+t)=\sum_1^{\infty}b_kf^{(k)}(t),\quad(5.11)$$

and from (5.10b) and the integral equation (4.19a) one sees that $b_k=c_k$. Then, multiplying (5.11) by $M(t'+t)\times[\zeta_1A(t)+\zeta_2A(t-\tau)]$ and integrating gives us an expression for the first term in the exponent of (5.10a):

$$\int_{-\infty}^{\infty}M^2(t'+t)[\zeta_1A(t)+\zeta_2A(t-\tau)]dt=\sum_1^{\infty}c_k^2.\quad(5.12)$$

The other term in (5.10a) may be found by a similar technique: Expansion of $\rho(s,t)$ gives

$$\rho(s,t)=\sum_1^\infty d_k f^{(k)}(t)=\sum_1^\infty \lambda_k f^{(k)}(s)f^{(k)}(t), \quad (5.13)$$

and multiplying by

$$M(t+t')M(s+t') \times[\zeta_1 A(t)+\zeta_2 A(t-\tau)][\zeta_1 A(s)+\zeta_2 A(s-\tau)]$$

and integrating over s,t yields

$$\int_{-\infty}^{\infty}\int \rho(s,t)M(t+t')M(s+t')\{\zeta_1^2 A(s)A(t) +\zeta_2^2 A(s-\tau)A(t-\tau)+2\zeta_1\zeta_2 A(s)A(t-\tau)\}dsdt =\sum_1^\infty \lambda_k c_k^2. \quad (5.14)$$

For the third term in (5.10a) we multiply (5.13) by $[\zeta_1 A(t)+\zeta_2 A(t-\tau)]\rho(x,\tau)$ and with the aid of the integral equation (4.19a) obtain

$$\int_{-\infty}^{\infty}[\zeta_1 A(t)+\zeta_2 A(t-\tau)]\rho(x,t)\rho(s,t)dt =\sum_1^\infty \lambda_k^2 f^{(k)}(x)f^{(k)}(s). \quad (5.15)$$

Multiplying this by the same expression used in deriving (5.14), replacing t by x, we obtain finally

$$\int_{-\infty}^{\infty}\int\int \{\zeta_1 A(x)+\zeta_2 A(x-\tau)\}\{\zeta_1 A(s)+\zeta_2 A(s-\tau)\} \times\{\zeta_1 A(t)+\zeta_2 A(t-\tau)\}M(t'+s)M(t'+x) \times\rho(s,t)\rho(x,t)dxdsdt=\sum_1^\infty \lambda_k^2 c_k^2 \quad (5.16)$$

in the general case.

Because of spatial limitations we now outline only the derivation of the leading term in W_2; the correction terms may be found in straightforward fashion with the help of (5.16) (expressed in frequency form), combined with (5.1), and then developed into a structure of the type (5.3), where, of course, $(W_2)_{N,0}$ is replaced by $(W_2)_{S+N,0}$ derived below. A further simplification (which does not in any way restrict the generality of the *method*) sets the signal equal to a constant: $M(t')=M$. Then (5.12) becomes

$$\sum_{k=1}^\infty c_k^2=M^2(\zeta_1+\zeta_2), \quad (5.17)$$

and the integrals occurring in (5.14) can be expressed in the frequency domain as

$$\int_{-\infty}^{\infty}\int A(s-t_1)A(t-\tau_2)\rho(s,t)dsdt =\int_{-\infty}^{\infty} H(f)|Y)i\omega)|^2 e^{i\omega(\tau_1-\tau_2)}df =H(0)\int_{-\infty}^{\infty}|Y(i\omega)|^2 e^{i\omega(\tau_1-\tau_2)} =H(0)\Phi_1(\tau_1-\tau_2)=H(0)\Phi_1(\tau_2-\tau_1), \quad (5.18)$$

since for the video narrow compared to the IF, $H(f)\doteq H(0)$. We get finally

$$\sum_{k=1}^\infty \lambda_k c_k^2=M^2H(0)\{(\zeta_1^2+\zeta_2^2)\Phi_1(0)+2\zeta_1\zeta_2\Phi_1(\tau)\}. \quad (5.19)$$

The characteristic function is the product of (5.10a) and (5.1), which for the normal term becomes on writing $H(0)=H_0$,

$$F_V(\zeta_1,\zeta_2;\tau)_{S+N,0} =\exp[i(\psi+M^2)(\zeta_1+\zeta_2)-(1/2)\{(\psi^2H_1+2\psi M^2H_0) \cdot((\zeta_1^2+\zeta_2^2)\Phi(0)+2\zeta_1\zeta_2\Phi_1(\tau))\}], \quad (5.20)$$

where we have made the substitution for $q(\tau)$:

$$q(\tau)=\Phi_1(\tau)\int_{-\infty}^{+\infty} H^2(f)df\equiv H_1\Phi_1(\tau), \quad (5.21)$$

which is valid for narrow videos. The corresponding probability density is

$$W_2(V_1,V_2;\tau)_{S+N,0}=[2\pi(\psi^2H_1+2\psi M^2H_0)\Phi_1(0)(1-\rho_0^2)^{\frac{1}{2}}]^{-1} \exp\left[\frac{[V_1-(\psi+M^2)]^2+[V_2-(\psi+M^2)]^2-2[V_1-(\psi+M^2)][V_2-(\psi+M^2)]\rho_0}{(\psi^2H_1+2\psi M^2H_0)\Phi_1(0)(1-\rho_0^2)}\right], \quad (5.22)$$

where ρ_0 is given by (5.5).

$$\left(\text{Note that } H_1=\int_{-\infty}^{\infty}\rho(\tau)^2dt=\int_{-\infty}^{\infty}H(f)^2df\right).$$

We observe in both (a) and (b) here, especially (5.3), (5.4), and (5.22) that in this approximation ($q(0)\neq 0$, $\Phi_1(0)\neq 0$) the probability densities appear to exist for V_1, $V_2<0$. Of course, this cannot strictly occur if $A(t)\geq 0$, $t>0$, since there can be no negative output of the final video filter in such cases; (see remarks at end of (a), Sec. 4). The higher-order terms in these approximations tend to reduce this contribution for V_1, $V_2<0$ to zero.

6. THE WIDE VIDEO FILTER

(a) *Narrow-Band Noise*

The case when the video band width is narrow compared to that of the IF stage has just been treated. We can also obtain expressions for the alternative limiting case where the video band width is large compared to the IF band width. The expected result should approach that of the distribution obtained with an infinitely wide, uniform video response.

If the IF band is limited (practically speaking) to a band** $2B$, and the video band width is b, one makes a Taylor series expansion

$$Y(i\omega)=1+\alpha_1\frac{i\omega}{2\pi B}+\alpha_2\left(\frac{\omega}{2\pi B}\right)^2+\cdots, \qquad \alpha_1\sim 0(B/b),\quad \alpha_2\sim 0(B/b)^2 \tag{6.1}$$

which certainly converges for $\omega<2\pi b$.

Evaluating the coefficients in the determinantal expansion of $F_V(\zeta_1\zeta_2\tau)$, see (4.15), we obtain

$$D_2(\tau)=\int_{-b}^{b}\int_{-b}^{b}H(f)H(f')\left[1+\frac{(\alpha_1^2+2\alpha_2)}{(2\pi)^2B^2}\times(\omega^2+2\omega\omega'+\omega'^2)\right]e^{i(\omega+\omega')\tau}df\,df'$$

$$+2\int_{b}^{\infty}\int_{b}^{\infty}H(f)H(f')\,|Y(i\omega+i\omega')|^2e^{i(\omega+\omega')\tau}df\,df, \tag{6.2}$$

retaining terms through the second order. The contribution of the second integral is negligible, since $H(f)$ falls off at least as $(1/\omega^2)$ and $Y(i\omega)$ at least as $(1/\omega)$ for $\omega>b/2\pi$. Also, if $H(f)$ falls off fast enough, the finite limits in the first integral may be replaced with negligible error by infinite limits. The latter step is not necessary, but is convenient for some cases.

Also we find [see (A.1) and (A3.9)] that

$$D_3(\tau,\tau)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}H(f'')H(f'+f'')H(f-f'')\cdot e^{-i\omega\tau}\left[1+(\alpha_1^2+2\alpha_2)\frac{(\omega^2+\omega'^2+\omega\omega')}{(2\pi)^2B^2}\right]df\,df'\,df'', \tag{6.3}$$

** This is true in many practical networks. In a stagger-tuned amplifier of n stages, for example, the normalized frequency response is

$$|Y(i\omega)|=\left[\frac{1}{1+(\omega/\omega_b)^n}\right]^{\frac{1}{2}}$$

The power, proportional to $|Y(i\omega)|^2$ is $\frac{1}{2}$ at $\omega=\omega_b$, and is $(1+2^{2n})^{-1}\doteq 2^{-2n}$ at $\omega=4\omega_b$. Thus for a 5-stage amplifier the ratio of power at $\omega=\omega_b$ to $\omega=4\omega_b$ is $512-1$.

which produces the desired approximation for the characteristic function

$$F_V(\zeta_1\zeta_2;\tau)_N\cong\left[F_V(\zeta_1,\zeta_2)_N|_\infty+\left(\frac{\alpha_1^2+2\alpha_2}{\beta^2}\right)\times\left\{\frac{q_1(0)\psi^2}{2}(\zeta_1^2+\zeta_2^2)+\psi^2\zeta_1\zeta_2q_1(\tau)+\frac{i\psi^3}{6}(\zeta_1^3+\zeta_2^3)q_1(0)+i\psi^3\left(q_1(\tau)+S'(\tau)-\frac{q_1(0)}{2}\right)\zeta_1^2\zeta_2+i\psi^3\left(q_1(\tau)+S'(-\tau)-\frac{q_1(0)}{2}\right)\zeta_1\zeta_2^2\right\}\right]^{-1}, \tag{6.4}$$

where

$$q_1(\tau)=\frac{-\partial^2[\rho^2(\tau)]}{(2\pi)^2\partial\tau^2}, \tag{6.5}$$

$$S'(\tau)=\left(\frac{1}{2\pi}\right)^2\int\int_{-\infty}^{\infty}\int\omega\omega'H(f'')H(f'+f'')\times H(f-f'')e^{-i\omega\tau}df\,df'\,df'', \tag{6.6}$$

$$\rho(\tau)=\int_{-\infty}^{\infty}H(f)e^{-i\omega\tau}df; \tag{6.7}$$

and $F_V(\zeta_1,\zeta_2;\tau)_N|_\infty$ is the characteristic function for infinite video when noise alone enters the quadratic detector. We get finally

$$F_V(\zeta_1\zeta_2;\tau)_N\cong F_V(\zeta_1\zeta_2;\tau)_N|_\infty\times\left[1-\frac{(2\alpha_2+\alpha_1^2)\psi^2}{B^2}F_V(\zeta_1\zeta_2;\tau)_N\Big|_\infty\theta(\zeta_1,\zeta_2)+\cdots\right], \tag{6.8}$$

where specifically

$$\theta(\zeta_1,\zeta_2)=\left[\frac{q_1(0)}{2}(\zeta_1^2+\zeta_2^2)+\zeta_1\zeta_2q_1(\tau)+\frac{i\psi}{6}(\zeta_1^3+\zeta_2^3)q_1(0)+i\psi\left[q_1(\tau)+S'(\tau)-\frac{q_1(0)}{2}\right]\zeta_1^2\zeta_2+i\psi\left[q_1(\tau)+S'(-\tau)-\frac{q_1(0)}{2}\right]\zeta_1\zeta_2^2\right]. \tag{6.9}$$

Thus

$$W_2(V_1,V_2;\tau)_N=W_2(V_1V_2;\tau)_{N,\infty}-\left(\frac{2\alpha_2+\alpha_1^2}{B^2}\right)\psi^2\times\theta\left(\frac{i\partial}{\partial V_1},\frac{i\partial}{\partial V_2}\right)W_2^{(1)}(V_1,V_2;\tau)_{N,\infty}, \tag{6.10}$$

where $W_2(V_1,V_2;\tau)_{N,\infty}$ is the distribution for the infinitely wide video, see (A2.16), and $W_2^{(1)}$ is obtained

from the convolution

$$W_2^{(1)}(V_1,V_2;\tau)_{N,\infty}=\int_{-\infty}^{\infty}\int W_2(z,y;\tau)_{N,\infty}\times W_2(V_1-z,V_2-y;\tau)_{N,\infty}dzdy. \quad (6.11)$$

(b) *Wide-Band Noise*

The results for wide-band noise follow easily from the work above. We find

$$F_V(\zeta_1\zeta_2;\tau)_N\cong F_V(\zeta_1\zeta_2;\tau)^{\frac{1}{2}}{}_{N,\infty}\left[1-\frac{1}{2}\left(\frac{2\alpha_2+{}_1\alpha^2}{B^2}\right)\times\psi^2F_V(\zeta_1,\zeta_2;\tau)_{N,\infty}\theta(\zeta_1,\zeta_2)+\cdots\right]. \quad (6.12)$$

Here $\theta(\zeta_1,\zeta_2)$ is given by (6.9) and the Fourier transform of $F_V(\zeta_1,\zeta_2;\tau)_{N,\infty}{}^{\frac{1}{2}}$ is the second-order probability density $W_2'(V_1,V_2,\tau)_N|_\infty$ obtained by the square-law rectification of *wide-band noise*. $W_2'(V_1,V_2;\tau)_N|_\infty$ can be obtained directly from the normal distribution, that is, from (A2.18). The result then is

$$W_2'(V_1,V_2;\tau)_N=\left[W_2'(W_1,V_2;\tau)_{N,\infty}-\frac{\psi^2}{2}\left(\frac{2\alpha_2+\alpha_1^2}{B^2}\right)\times\theta\left(i\frac{\partial}{\partial x_1},i\frac{\partial}{\partial x_2}\right)W_2^{(1)}(x_1x_2;\tau)_N\right], \quad (6.13)$$

where $W_2^{(1)}(V_1,V_2,\tau)_N$ is given here by

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}W_2(z,y;\tau)_{N,\infty}W_2'(x_1-z,\,x_2-y;\tau)_{N,\infty}dzdy. \quad (6.14)$$

(c) *First-Order Approximation (Narrow-Band Noise)*

As an example, we observe that it is possible to get approximations to $W_1(V)_N$ for narrow-band noise by solving the integral equation for the first-order case, *viz.*,

$$\int_{-\infty}^{\infty}Y(i\omega_1-i\omega_2)H(f_1)g^{(k)}(f_2)df_2=\lambda_kg^{(k)}(f). \quad (A3.11)$$

The present method differs from that used by Kac and Siegert in that it yields any number of eigenvalues, depending on how closely one approximates the video filter by a Taylor series. Certain limitations on the types of filter for which approximations to $W_1(V)$ can be found by the method of Kac and Siegert are removed.

Using the Taylor series expansion (6.1) and substituting into the integral equation above, we obtain

$$H(f_1)\int_{-b}^{b}\left[1+i\alpha\frac{(\omega_1-\omega_2)}{2\pi B}+\frac{\alpha_2(\omega_1-\omega_2)^2}{(2\pi B)^2}\right]g^{(k)}(f_2)df_2+H(f)\int_b^{\infty}Y(i\omega_1-i\omega_2)g^{(k)}(f_2)df_2+H(f_1)\int_{-\infty}^{b}Y(i\omega-i\omega_2)g^{(k)}(f_2)df_2=\lambda_kg^{(k)}(f). \quad (6.15)$$

Expanding and neglecting the last two integrals, we find that††

$$H(f_1)\left[\int_{-\infty}^{\infty}g^{(k)}(f_2)\left[1+i\frac{\alpha_1\omega_1}{2\pi B}+\frac{\alpha_2\omega_1{}^2}{(2\pi B)^2}\right]df_2+\int_{-\infty}^{\infty}\omega_2g^{(k)}(f_2)\left[-\frac{i\alpha_1}{2\pi B}-\frac{2\alpha_2\omega_1}{(2\pi B)^2}\right]df_2+\int_{-\infty}^{\infty}\omega_2{}^2g^{(k)}(f_2)\frac{\alpha_2}{(2\pi B)^2}df_2\right]=\lambda_kg^{(k)}(f). \quad (6.16)$$

Solving the integral equation, we obtain finally these important eigenvalues:

$$\lambda_1\doteq 1+\left(\frac{2\alpha_2+\alpha_1{}^2}{B^2}\right)\int_{-\infty}^{\infty}f^2H(f)df, \quad (6.17a)$$

which is the same result for λ_1 as obtained by Kac and Siegert.‡‡ We have also

$$\lambda_2\doteq\left(\frac{2\alpha_2+\alpha_1{}^2}{B^2}\right)\frac{\rho''(0)}{(2\pi)^2} \quad (6.17b)$$

and

$$\lambda_3=\frac{2\alpha_2{}^3}{(2\alpha_2+\alpha_1{}^2)B^4}\left[\int_{-\infty}^{\infty}f^2H(f)df-\int_{-\infty}^{\infty}f^4H(f)df\right]. \quad (6.17c)$$

Using Eqs. (4.46), (4.47) of reference 5 and the foregoing, we may write approximately

$$W_1(V)_N=\left[\frac{1-2\alpha}{\psi}\right]\exp[-V(1-\alpha)/\psi]-\frac{1}{\psi}\exp\frac{V}{\alpha\psi}-\frac{\gamma}{\alpha\psi}\exp(-V/\gamma\psi),\quad V<0, \quad (6.18)$$

where $1-\lambda_1=\lambda_2=-\alpha$, $\lambda_3=\gamma$; ($\alpha>0$, $\gamma<0$) and all terms of $0(B/b)^2$ are included. In this case $V>|\alpha|\psi$, as $\alpha\to 0$, and we obtain on noting that $|\gamma|\ll(\alpha)$ the further

†† The infinite limits apply in the same sense as the cases considered in Eqs. (6.2) and (6.3).

‡‡ Note that for certain functions $H(f)$ we cannot integrate between infinite limits, but must use $-b$ and $+b$.

approximation

$$W(V)_N \doteq \frac{e^{-V/\psi}}{\psi}\left[1-2\alpha+\alpha\frac{V}{\psi}\right], \quad (V>0). \quad (6.19)$$

(d) A First-Order Approximation. Narrow-Band Signal and Noise

A further example in the case of wide video filters is provided when the input is the sum of signal and noise; the main features of the approximation are outlined below.

We start with (4.16) and the integral equation (4.17). Our approximation is limited to the calculation of a single eigenvalue (λ_1), and a single eigenfunction $f^{(1)}(t)$. Calling

$$\left(\int_{-\infty}^{\infty} M(t+t')A(t)f^{(1)}(t)dt\right)^2 = k^2/\zeta_1,$$

and writing $\lambda_1 = 1+\alpha$, *viz.* (6.18), we see that Eq. (4.16) becomes

$$F_V(\zeta)_{S+N} \doteq (1-i\psi\zeta_1-\alpha i\zeta_1\psi)^{-1} \times \exp[i\zeta_1\psi k^2/(1-i\zeta_1\psi-\alpha i\zeta_1\psi)]. \quad (6.20)$$

Making the approximation for small α

$$(1-i\psi\zeta_1-\alpha i\zeta_1\psi)^{-1} = (1-i\zeta\psi)^{-1}+\alpha i\zeta_1\psi/(1-i\zeta_1\psi)^2$$

and expanding the corresponding term (containing α) in the exponent of (6.20), we obtain finally

$$F_V(\zeta_1)_{S+N} = \frac{1}{(1-i\zeta_1\psi)}\exp\left(\frac{i\zeta_1\psi k^2}{1-i\zeta_1\psi}\right) \times [1+i\alpha\zeta_1\psi F_V(\zeta)_{S+N,\infty}-\alpha\psi^2\zeta_1^2k^2F_V^2(\zeta)_{S+N,\infty}], \quad (6.21)$$

where we have retained terms through $0(\alpha)$ only. The Fourier transform of (6.21) is easily found, see (A2.1), (A2.2), and we can write finally

$$W_1(V)_{S+N} \backsimeq W_1(V)_{S+N}|_\infty - \alpha\psi\frac{\partial}{\partial V_1}W_1^{(1)}(V)_{S+N} + k^2\alpha\psi\frac{\partial^2}{\partial V_1^2}W^{(2)}(V)_{S+N} \quad (6.22)$$

to this order of approximation; $W_1(V)_{S+N}|_\infty$ is given by (A2.2), and

$$W_1^{(1)}(V)_{S+N} = \int_{-\infty}^{\infty} e^{-x/\psi}W_1(V-x)_{S+N}|_\infty dx/\psi,$$

$$W_1^{(2)}(V)_{S+N} = \int_{-\infty}^{\infty} e^{-x/\psi}W_1^{(1)}(V-x)_{S+N}dx/\psi.$$

7. A SPECIAL VIDEO FILTER

If the IF noise and signal wave are confined to a band $2f_b$, this implies that the slowly varying part has a maximum frequency spread f_b. A square-law detector doubles this spread to $2f_b$, so that the maximum band that the detected wave occupies is $2f_b$. Thus, if the video filter is flat over this band, one would expect the same result as for an infinitely wide filter. The expected result can be shown as follows:

We choose

$$Y(i\omega) = e^{-i\omega}, \quad 0<|f|<2f_b,$$

$$H(f) = 0\begin{cases} f<-f_b \\ f>f_b \end{cases}. \quad (7.1)$$

The condition $H(f)=0$ at frequencies greater than f_b from the center of the IF pass band is often practically met, as for example in a stagger-tuned amplifier with many stages.

Using the frequency form of the integral equation (A3.11), choosing

$$g(f) = [k_1\zeta_1H(f)+k_2\zeta_2e^{-i\omega t}H(f)]e^{i\omega} \quad (7.2)$$

and substituting this into the integral equation, we obtain finally

$$[\zeta_1\{k_1\zeta_1+k_2\zeta_2\rho(\tau)\}+\zeta_2e^{-i\omega\tau}\{k_1\zeta_1\rho(\tau)+k_2\zeta_2\}] = \lambda(k_1\zeta_1+k_2\zeta_2e^{-i\omega\tau}). \quad (7.3)$$

This relation can be satisfied if we set

$$(\zeta_1-\lambda)k_1+\zeta_2k_2\rho(\tau)=0, \quad (7.4a)$$

$$\zeta_1k_1\rho(\tau)+k_2(\zeta_2-\lambda)=0, \quad (7.4b)$$

which leads to the same values of λ as found in Appendix II (A2.9), for the case of an infinitely wide video.

8. CONCLUDING REMARKS

The preceding analysis has shown that the second- and higher-order distribution functions for noise or signal and noise which have been rectified by a quadratic detector and then filtered can be represented only in closed form in terms of the characteristic functions of these distributions. Even so, such representations imply the solution of a difficult integral equation.

In many instances of practical importance, we are interested in limiting cases, i.e., either a very wide or a very narrow video. For wide video, by expanding the characteristic function we can obtain approximations to the second-order distribution function, with correction terms obtained from the known distribution function for infinitely wide video response. As seen in Sec. 7, a wide video means a video band pass that is of the order of the band width at the IF frequency which contains most of the noise power.

Again, for cases where another nonlinear operation follows video (or audio) filtering, the characteristic function is of value and the approximations developed here can be used to advantage. As the video filter becomes more narrow, terms in the denominator of

$F_V(\zeta_1,\zeta_2;\tau)$ representing powers of (ζ_1,ζ_2) greater than unity, assume more importance. It appears that powers of ζ up to the third yield results which imply a first-order correction to the distribution for the infinitely wide post detection filter.

As may be seen from the approximations developed for $W_1(x)$ in Sec. 6, more terms in the Taylor series approximation to the video filter function must be considered as the video becomes narrower, and more eigenvalues of significance are obtained.

In the extreme of a narrow-video filter one would expect from physical considerations that the distribution becomes normal. The results of the analytical work show to what extent the actual distribution departs from the normal for videos narrow compared to the IF band. Correction terms have coefficients whose magnitudes are of the order of powers of the ratio

$$\left(\frac{\text{video width}=b}{IF\text{ width}=2B}\right).$$

APPENDIX I. MATRIX EQUATIONS

In Eq. (2.8) the exponent

$$\sum_{k,l=1}^{n}\left(\frac{\alpha^{kl}}{|\alpha|\psi}-i\zeta_k\delta_{kl}\right)x_kx_l$$

is represented in matrix form by $\mathbf{x}'\mathbf{B}\mathbf{x}$, where $\mathbf{x}$ is a column vector, $\mathbf{x}'$ the transpose, and $\mathbf{B}$ is a symmetrical matrix with components $B_{kl}=(\alpha^{kl}/\psi|\alpha|)-i\zeta_k\delta_{kl}$; $|\alpha|$ is the determinant of $\boldsymbol{\alpha}$ and α^{kl} the cofactor of α_{kl} in $\boldsymbol{\alpha}$. An orthogonal matrix can be found such that $\mathbf{P}'\mathbf{B}\mathbf{P}=\boldsymbol{\Lambda}\mathbf{I}$, where $\boldsymbol{\Lambda}$ is diagonal ($\mathbf{I}=\delta_{kl}$). Letting $\mathbf{Z}=\mathbf{P}^{-1}\mathbf{x}$, $\mathbf{x}'\mathbf{B}\mathbf{x}$ becomes $\sum_{j=1}^{n}\lambda_jZ_j^2$, where the λ_j are the components of $\boldsymbol{\Lambda}=[\lambda_j\delta_{jk}]$.

By a simple quadrature one obtains

$$\phi_N(\zeta_1,\cdots\zeta_n)=\psi^{-n}|\alpha|^{-1}\prod_{j=1}^{n}\lambda_j^{-1}=\psi^{-n}|\mathbf{B}\boldsymbol{\alpha}|^{-1}, \quad \text{(A1.1)}$$

since $\prod_{j=1}^{n}\lambda_j=|\boldsymbol{\Lambda}|=|\mathbf{B}|$. This gives (2.9) with a little further manipulation, since

$$\mathbf{B}=\boldsymbol{\alpha}^{-1}\psi^{-1}-i\boldsymbol{\zeta}; \quad \boldsymbol{\alpha}^{-1}=[\alpha^{kl}/|\alpha|]=\boldsymbol{\varrho}^{-1}. \quad \text{(A1.2)}$$

Now, by finding a symmetrical matrix $\mathbf{Q}$, such that $\mathbf{Q}(\boldsymbol{\Lambda}-i\mathbf{I})\mathbf{Q}^{-1}=\boldsymbol{\Lambda}'\mathbf{I}$, where

$$[\boldsymbol{\Lambda}]_{k,j}=\zeta_k\psi\rho_{k,j}+i\delta_{k,j}, \quad \text{(A1.3)}$$

inasmuch as $\phi_N(\zeta_1,\cdots\zeta_n)=i^n|\boldsymbol{\Lambda}|^{-1}=|\mathbf{I}-i\boldsymbol{\Lambda}'|^{-1}$, we obtain the result (2.10a), *viz.*,

$$\phi_N(\zeta_1,\cdots\zeta_n)=\prod_{j=1}^{n}(1-\lambda_j')^{-1}. \quad \text{(A1.4)}$$

The equations for the rows of $\mathbf{Q}$ become the equations for the eigenvectors $\mathbf{f}_{(j)}(\boldsymbol{\Lambda}-i\mathbf{I})=\boldsymbol{\lambda}'\mathbf{f}_{(j)}$, or

$$\psi\sum_{k}^{n}\zeta_k\rho_{km}f_{jk}=\lambda_j'f_{jm}, \quad \text{(A1.5)}$$

which is Eq. (2.10b).

APPENDIX II. DISTRIBUTIONS FOR THE INFINITELY WIDE VIDEO FILTER

When the video filter is infinitely wide we expect the general results of Sec. 4 to reduce to the more familiar distributions of the amplitude following square-law rectification alone. The latter may, of course, be obtained directly from (2.10)–(2.12) for noise alone, or from (3.5a), (3.7) for a signal and noise, or, as was originally done in first-order cases, by a direct transformation of variables $V_0=\beta^2V_{in}^2$ according to the detector's law, applied to the original normal distribution for the noise (see Sec. 2). Accordingly, we now verify that our general results for arbitrary filtering, in the case of the first- and second-order densities, reduce to the expected forms when the video filter is allowed to become indefinitely wide.§§ The first-order densities for signal and noise follow at once.

(I) First-Order Densities

For an infinitely wide video, the filter's weighting function is $A(t)=\delta(t-0)$, and the solution to (4.17) is at once $f(s)=C\rho(s)$ and $\lambda_1'=1$. The normalization condition gives $C=\zeta_1^{-1/2}$, whereupon (4.16) becomes

(*a*) *Narrow-Band Waves.*—

$$F_V(\zeta_1)_{S+N}|_\infty=\phi_{S+N}(\zeta_1)=(1-i\zeta_1\psi)^{-1}\times\exp[i\zeta_1M(t')^2/(1-\psi\zeta_1)]. \quad \text{(A2.1)}$$

The corresponding probability density is found with the help of Pair No. 650.0 [G. A. Campbell and R. M. Foster, *Fourier Integrals for Practical Applications* (D. Van Nostrand Company, Inc., New York, 1947)] to be

$$W_1(V)_{S+N}|_\infty=\psi^{-1}\exp(-(V+M^2)/\psi)I_0(2MV^{\frac{1}{2}}/\psi), \quad V>0;\ =0;\ V<0;\ (\psi=2\langle V_N^2\rangle \quad \text{(A2.2)}$$

which is a well-known result.[5,14] For noise alone one has

$$W_1(V)|_\infty=\psi^{-1}e^{-V/\psi}, \quad V>0;\ =0;\ V<0, \quad \text{(A2.3)}$$

which is a χ^2 density with 2 degrees of freedom, as expected. [From (2.6a) we recall that since $\beta^2=2$ here, $\psi=2\langle V_N^2\rangle$, where $\langle V_N^2\rangle$ is the mean-square noise output of the IF filter.]

(*b*) *Broad-band Waves.*—

Here (A2.1) is appropriately modified, to give us

$$W_1(V)_{S+N}|_\infty=\int_{-\infty}^{\infty}\exp[(-i\zeta_1V+i\zeta_1s^2)/(1-i2\psi\zeta_1)]\times\left(\frac{d\zeta_1}{(1-\zeta_1{}^2\psi)^{\frac{1}{2}}2\pi}\right)$$
$$=(2\pi V\psi)^{-\frac{1}{2}}\exp[-(s^2+V)/2\psi]\cosh[V^{\frac{1}{2}}s/\psi], \quad V>0;\ =0;\ V<0, \quad \text{(A2.4)}$$

where we have again used transform pair No. 650.0. This result has also been obtained by Emerson[15] in a recent article. For noise alone we get a χ^2 distribution with 1 degree of freedom:

$$W_1(V)_N|_\infty=(2\pi V_\psi)^{-\frac{1}{2}}e^{-V/2\psi}, \quad V>0;\ =0;\ V<0,\ (\psi=\langle V_N^2\rangle). \quad \text{(A2.5)}$$

(II) Second-Order Densities

Here the calculations are more involved. From (4.18) we have for general post-detection filtering

$$F_V(\zeta_1,\zeta_2;\tau)_{S+N}=\prod_{k=1}^{\infty}(1-i\lambda_k)^{-\alpha}\exp\left\{\left(i\int_{-\infty}^{\infty}M(t'+t)\times[\zeta_1A(t)+\zeta_2A(t-\tau)]f^k(t)dt\right)^2\Big/(1-i\lambda_k)\right\}, \quad \text{(4.18)}$$

where $\alpha=1$ for narrow-band waves, and $\alpha=\frac{1}{2}$ for broad-band disturbances, with M replaced by S. The associated integral

§§ The carrier or central-frequency zone in the narrow-band case is, however, not passed by this "infinitely" wide video.

[14] S. A. Goudsmit, RL 43-21; J. C. Slater, RL V-23; D. O. North, RCA Tech. Rept. PTR6C; K. A. Norton and V. D. Landon, Proc. Inst. Radio Engrs. 30, 425 (1942). (RL = Radiation Laboratory, MIT.)

[15] R. C. Emerson, J. Appl. Phys. 24, 1168 (1953). (Here $V=\psi Y$, $S^2=\psi X$, and our $\psi=\langle V_N^2\rangle$, since here $\beta^2=1$.) In general, for arbitrary β we must replace ψ by $\beta^2\psi/2=\beta^2\langle V_N^2\rangle/2$ for narrow-band waves, and $\beta^2\psi=\beta^2\langle V_N^2\rangle$ for broad-band waves, in all results of the present paper.

equation is

$$\psi\int_{-\infty}^{\infty}[\zeta_1 A(t)+\zeta_2 A(t-\tau)]\rho(t,s)f^{(k)}(t)dt=\lambda_k f^{(k)}(s) \quad \text{(A2.6)}$$

and since $A(t)=\delta(t-0)$; $\rho(0,s)=\rho(-s)=\rho(s)$; $\rho(\tau,-s)=\rho(\tau-s)$, etc., Eq. (A2.6) becomes

$$\zeta_1\rho(s)f^{(k)}(0)+\zeta_2\rho(\tau-s)f^{(k)}(\tau)=\lambda_k f^{(k)}(s)\psi^{-1}, \quad \text{(A2.7a)}$$

subject to the normalization conditions

$$\zeta_1 f^{(k)}(0)^2+\zeta_2 f^{(k)}(\tau^2)=1. \quad \text{(A2.7b)}$$

In general, the integral equation (A2.6) is here satisfied by

$$f^{(k)}(s)=C_1^{(k)}(\zeta_1,\zeta_2)\zeta_1\rho(s)+C_2^{(k)}(\zeta_1,\zeta_2)\zeta_2\rho(s-\tau). \quad \text{(A2.8)}$$

Substituting (A2.8) into (A2.7a) (or (A2.6)), rearranging terms, and setting the coefficients of $\rho(s)$ and $\rho(s-\tau)$ equal to zero, since the equations are then satisfied by arbitrary $\rho(s)$, arbitrary s, τ, we obtain

$$C_1^{(k)}\left(\frac{\lambda_k}{\psi}-\zeta_1\right)-C_2^{(k)}\zeta_2\rho(\tau)=0;$$

$$-C_1^{(k)}\zeta_1\rho(\tau)+C_2^{(k)}\left(\frac{\lambda_k}{\psi}-\zeta_2\right)=0. \quad \text{(A2.9)}$$

For nontrivial values of C_1, C_2 we determine from this the eigenvalues λ_k, by setting the determinant of the coefficients C_1, C_2 equal to zero, obtaining two values

$$\lambda_k=\frac{\psi}{2}\{\zeta_1+\zeta_2\pm[\zeta_1^2+\zeta_2^2-4\zeta_1\zeta_2(1-\rho^2)]^{\frac{1}{2}}\}; \quad k=1,2 \quad \text{(A2.10)}$$

and hence *four* coefficients $C_1^{(k)}$, $C_2^{(k)}$. From the two relations (A2.8) in (A2, 7a), and the two equations (A2.7b) for normalization we get the necessary number of expressions to determine the coefficients, and hence the eigenfunctions. Unlike the case of noise alone, the eigenfunctions, as well as the eigenvalues λ_k ($k=1, 2$) are needed when there is a signal, as can be seen from (4.18) above. The desired probability density $W_2(V_1,V_2,\tau)_{S+N}|_\infty$ then follows by inversion, in the usual way.

Since $C_1^{(k)}$, $C_2^{(k)}$ are, however, complicated functions of ζ_1, and ζ_2, the actual inversion is not easy. The expected result may be obtained, none the less, by a direct transformation of W_2, where $V_{\text{out}}=\beta^2 V_{\text{in}}{}^2=R_{\text{in}}{}^2$. For narrow-band cases, using a result of Middleton,[16] we have at once

(*narrow-band*):

$$W_2(V_1,V_2;\tau)_{S+N}|_\infty=\{[\psi^2(1-\rho^2)]^{-1}\exp[-(V_1+V_2)/\psi(1-\rho^2)]$$
$$\times\exp[-(M_1{}^2+M_2{}^2-2M_1M_2\rho)/\psi(1-\rho^2)]\}$$
$$\cdot\sum_{M=0}^{\infty}\epsilon_m I_m\left(\frac{2\rho(V_1V_2)^{\frac{1}{2}}}{\psi(1-\rho^2)}\right)I_m\left(\frac{M_1-M_2\rho}{\psi(1-\rho^2)}2V_1{}^{\frac{1}{2}}\right)I_m\left(\frac{M_2-M_1\rho}{\psi(1-\rho^2)}2V_2{}^{\frac{1}{2}}\right),$$
$$V_1, V_2>0;\ V_1, V_2<0 \quad \text{(A2.11)}$$

and where $M_1=M(t_1)$; $M_2=M(t_2)$.

In the broad-band situation a direct transformation of (2.9), using (2.10), of reference (16) and the fact that $V_{\text{out}}=\beta^2 V_{\text{in}}{}^2$ gives us immediately ($\beta^2=1$):

(*broad-band*):

$$W_2(V_1,V_2;\tau)_{S+N}|_\infty=(\{2\pi\psi[V_1V_2(1-\rho^2)]^{\frac{1}{2}}\}^{-1}$$
$$\cdot\exp\{-[(V_1+V_2-2\rho(V_1V_2)^{\frac{1}{2}}]/2\psi(1-\rho^2)\})$$
$$\cdot\exp\{-[S_1{}^2+S_2{}^2-2\rho S_1S_2-2(V_1)^{\frac{1}{2}}(S_1-\rho S_2)$$
$$-2(V_2)^{\frac{1}{2}}(S_2-\rho S_1)]/2\psi(1-\rho^2)\},$$
$$V_1; V_2>0;\ =0, V_1, V_2<0. \quad \text{(A2.12)}$$

Where there is no signal, we do not need the eigenfunctions, only the eigenvalues (A2.10). The characteristic function (4.18) reduces now in the narrow-band case to

$$F_V(\zeta_1,\zeta_2;\tau)_N|_\infty=\phi(\zeta_1,\zeta_2)_N=[1-\psi^2\lambda_1\lambda_2-i\psi(\lambda_1+\lambda_2)]^{-1}$$
$$=[1-i\psi(\zeta_1+\zeta_2)-\psi^2\zeta_1\zeta_2(1-\rho^2(\tau))]^{-1}. \quad \text{(A2.13)}$$

[16] D. Middleton, Quart. Appl. Math. **5**, 445 (1948), Eq. (5.17).

The second-order density is obtained from the Fourier transform of (A2.13); letting $1-\rho^2=k$, we can write this transform as

$$W_2(V_1,V_2;\tau)_N|_\infty=\psi^{-1}(2\pi)^{-2}\int_{-\infty}^{\infty}\frac{d\zeta_2 e^{-i\zeta_2V_2}}{(1-i\psi k\zeta_2)}\int_{-\infty}^{\infty}d\zeta_1 e^{-i\zeta_1V_1}$$
$$\times\left[\zeta_1+\frac{i}{\psi}\left(\frac{1-i\psi\zeta_2}{1-i\psi\zeta_2 k}\right)\right]^{-1}. \quad \text{(A2.14)}$$

The second integral in (A2.14) has a pole in the lower plane for all real values of ζ_2. For $V_1<0$ the integral is zero, and since (A2.14) is symmetrical in ζ_1V_1, ζ_2V_2, it vanishes also for $V_2<0$. For $V_1>0$, we get for the integral in ζ_1 the value

$$-2\pi i\exp[-V_1(1-i\psi\zeta_2)/\psi(1-i\psi k\zeta_2)].$$

Letting $\mu=k\psi\zeta_2+i=[(V_1/V_2)(1+k)]^{\frac{1}{2}}\eta$, we have

$$W_2=\frac{ie^{-(V_1+V_2)/k\psi}}{2\pi k\psi^2}\int_{-\infty+i}^{\infty+i}[\exp-i(V_1V_2(1+k)/k^2\psi^2)^{\frac{1}{2}}]$$
$$\times[\eta-\eta^{-1}]d\eta/\eta. \quad \text{(A2.15)}$$

This integral is recognized as a modified Bessel function of zeroth order, giving us finally the result

$$W_2(V_1,V_2;\tau)_N|_\infty=[\psi(1-\rho^2)]^{-1}\exp[-(V_1+V_2)/\psi(1-\rho^2)]$$
$$\times I_0\left(\frac{2\rho(V_1V_2)^{\frac{1}{2}}}{\psi(1-\rho^2)}\right),\ V_1, V_2>0;\ =0,\ V_1,V_2<0 \quad \text{(A2.16)}$$

for the narrow-band case, which is immediately verified from (A2.11) on setting $M_1=M_2=0$.

The broad-band case may be obtained in a similar way, where now the characteristic function is

$$F_V(\zeta_1,\zeta_2;\tau)_N|_\infty=[1-i\psi(\zeta_1+\zeta_2)-\psi^2\zeta_1\zeta_2(1-\rho^2)]^{-\frac{1}{2}}. \quad \text{(A2.17)}$$

We have finally

(*broad-band*):

$$W_2(V_1,V_2;\tau)_N|_\infty=\{2\pi\psi[V_1V_2(1-\rho^2)]^{\frac{1}{2}}\}^{-1}$$
$$\times\exp\{-[V_1+V_2-2\rho(V_1V_2)^{\frac{1}{2}}]/2\psi(1-\rho^2)\},$$
$$V_1, V_2>0;\ =0, V_1,V_2<0, \quad \text{(A2.18)}$$

as expected, see (A2.12) when $S_1=S_2=0$.

APPENDIX III. EXPANSION OF THE CHARACTERISTIC FUNCTION $F_v(\zeta)$. TIME AND FREQUENCY REPRESENTATIONS

As we have seen, a form of the characteristic function which is useful in the important limiting cases of wide or narrow video filters is given by the expansion of the reciprocal of the Fredholm determinant associated with the integral equation (4.10b) in the general case, and with (4.13b) for $n=2$. For the latter, the characteristic function is Δ^{-1}, see (4.15).

To represent the expansion of $F_V(\zeta_1,\zeta_2;\tau)_N$ compactly in both the time and frequency domains, the following Fourier transforms are needed, relating system-, weighting-, correlation functions, and intensity spectra associated with the IF and video filters. We have for the IF filter

$$\rho(\tau)=\int_{-\infty}^{\infty}H(f)e^{i\omega\tau}df; \quad \rho(0)=1 \quad \text{(A3.1)}$$

$$H(f)=\int_{-\infty}^{\infty}\rho(\tau)e^{-i\omega\tau}d\tau=|B(f\pm f_0)|^2, \quad \text{(A3.2)}$$

$$\int_{-\infty}^{\infty}H(f)df=1; \quad \int_{-\infty}^{\infty}H^2(f)df=\int_{-\infty}^{\infty}\rho^2(\tau)d\tau\equiv H_1. \quad \text{(A3.3)}$$

For the video filter we write

$$Y(i\omega)=\int_{-\infty}^{\infty}A(t)e^{-i\omega t}dt; \quad A(t)=\int_{-\infty}^{\infty}Y(i\omega)e^{i\omega t}df;$$
$$t<0,\ =0 \quad \text{(A3.4)}$$

$$|Y(i\omega)|^2=\int_{-\infty}^{\infty}\Phi_1(t)e^{-i\omega t}dt; \quad \Phi_1(t)=\int_{-\infty}^{\infty}|Y(i\omega)|^2e^{i\omega t}df \quad \text{(A3.5)}$$

$$=\int_{-\infty}^{\infty}A(t')A(t'\pm t)dt'=\Phi_1(\pm t), \quad \text{(A3.5a)}$$

and

$$\Phi_2(\tau_1,\tau_2)\equiv\int_{-\infty}^{\infty}A(t')A(t'+\tau_1)A(t'+\tau_2)dt'$$
$$=\int\int_{-\infty}^{\infty}Y(i\omega)Y(i\omega')Y(i\omega+i\omega')^*\exp[i\omega'(\tau_1-\tau_2)-i\omega\tau_2]d\,fd\,f'. \quad (A3.6)$$

When the determinantal expansion is carried out through terms of order 3 in (ζ_1,ζ_2), the result is finally

$$F_V(\zeta_1,\zeta_2;\tau)_N=\Big\{1-i\psi(\zeta_1+\zeta_2)$$
$$-\frac{\psi^2}{2!}[(\zeta_1^2+\zeta_2^2)(1-D_2(0))+2\zeta_1\zeta_2(1-D_2(-\tau))]$$
$$+\frac{i\psi^3}{3!}[(\zeta_1+\zeta_2)^3+3(\zeta_1+\zeta_2)\{(\zeta_1^2+\zeta_2^2)D_2(0)+2\zeta_1\zeta_2D_2(-\tau)\}$$
$$+2(\zeta_1^3+\zeta_2^3)D_3(0,0)+2\zeta_1^2\zeta_2\{2D_3(-\tau,0)+D_3(\tau,\tau)\}$$
$$+2\zeta_1\zeta_2^2\{2D_3(\tau,0)+D_3(-\tau,-\tau)\}]+\cdots\Big\}^{-1}, \quad (A3.7)$$

where for broad-band noise $\{\ \}^{-1}$ is replaced by $\{\ \}^{-\frac{1}{2}}$, as usual. The functions D_2, D_3 are specifically

$$D_2(\tau)=\int_{-\infty}^{\infty}\rho^2(t)\Phi_1(t+\tau)dt=\int\int_{-\infty}^{\infty}H(f)H(f')$$
$$\times|Y(i\omega+i\omega')|^2\exp[i(\omega+\omega')\tau]d\,fd\,f', \quad (A3.8)$$

$$D_3(\tau,\tau)=\int\int_{-\infty}^{\infty}\Phi_2(u+\tau,v+\tau)\rho(u)\rho(v)\rho(u-v)dudv$$
$$=\int\int\int_{-\infty}^{\infty}Y(i\omega)Y(i\omega')Y(i\omega+i\omega')^*H(f-f'')$$
$$\times H(f'+f'')H(f'')e^{-i\omega\tau}d\,fd\,f'd\,f''. \quad (A3.9)$$

Further relations of some use are

$$2D_3(-\tau,0)+D_3(\tau,\tau)=3\int\int\int_{-\infty}^{\infty}Y(i\omega)Y(i\omega')Y(i\omega+i\omega')^*$$
$$\times H(f-f'')H(f'+f'')H(f'')e^{-i\omega\tau}d\,fd\,f'd\,f'' \quad (A3.10)$$

and $2D_3(\tau,0)+D_3(-\tau,-\tau)$ is precisely (A3.10) with ω replaced by ω' in the exponent. The integral equation (4.13b) for $n=2$ is transformed in a similar way to

$$\psi H(f)\int_{-\infty}^{\infty}\{\zeta_1+\zeta_2\exp[i(\omega'-\omega)\tau]\}Y(i\omega-i\omega')g^{(k)}(f')d\,f'$$
$$=\lambda_k g^{(k)}(f), \quad (A3.11)$$

where it is assumed that the $f^{(k)}(t)$ possess a transform

$$f^{(k)}(t)=\int_{-\infty}^{\infty}g^{(k)}(f)e^{i\omega t}d\,f;\quad g^{(k)}(f)=\int_{-\infty}^{\infty}f^{(k)}(t)e^{-i\omega t}dt. \quad (A3.12)$$

APPENDIX IV. REMARK ON THE CALCULATION OF MOMENTS

The moments of the distribution for $W_2(V_1,V_2;\tau)$ can be obtained easily from the Taylor series expansion for $F_V(\zeta_1,\zeta_2;\tau)$, as is well known. That is, we can write, when these quantities exist,

$$\langle V_1^kV_2^l\rangle=A_{kl}(-i)^{k+l}/l!k!,$$

where the A_{kl} are the coefficients of the series expansion. One then uses the characteristic function expansion of $F_V(\zeta_1,\zeta_2;\tau)$ to obtain the corresponding coefficients in the Taylor's series.

One result of this calculation is

$$\langle V_1V_2\rangle=\psi^2\Big(1+\int\int_{-\infty}^{\infty}H(f)H(f')|Y(f+f')|^2$$
$$\times\exp[-i(\omega+\omega')\tau]d\,fd\,f'\Big), \quad (A4.1)$$

the correlation function of the output wave.

Letters to the Editor

Diffraction of Electromagnetic Waves by an Aperture in a Large Screen

L. LEWIN
Standard Telecommunication Laboratories Limited, Enfield, Middlesex, England
(Received April 5, 1954)

IN a letter in your February issue, J. H. Crysdale[1] deals with Bekefi's[2] analysis of the field arising from a magnetic Hertzian vector Π^*. In attempting to satisfy boundary conditions with a single component Π_y^*, the relation

$$\Pi_y^* = \Pi_{iy}^* \tag{1}$$

in the aperture was stated to be consistent with the boundary conditions on the magnetic fields derived from

$$H_x = \partial^2\Pi_y^*/\partial x\partial y$$
$$H_y = \partial^2\Pi_y^*/\partial y^2 + k^2\Pi_y^*. \tag{2}$$

Obviously the equality of Π_y^* in the two regions ensures the equality of H_x and H_y. However, the fields calculated from (1) were not in agreement with experiment.

Although consistent with (2), Eq. (1) is not the most general solution arising from (2). Continuity of H_y leads to

$$\partial^2\Pi_y^*/\partial y^2 + k^2\Pi_y^* = \partial^2\Pi_{iy}^*/\partial y^2 + k^2\Pi_{iy}^* \tag{3}$$

or

$$\Pi_y^* = \Pi_{iy}^* + A\cos ky + B\sin ky, \tag{4}$$

where A and B are independent of y.

Continuity of H_x requires further that A and B be constant. Only if they are set equal to zero do we get Eq. (1).

Using (4) instead of (1) the solution has been obtained to two problems which satisfy the condition imposed by the choice of a single component for Π^*, namely, the vanishing of E_y. These are (a) the radiation of a rectangular wave-guide aperture in an infinite conducting screen[3] and (b) the transmission through a symmetrically situated rectangular window in a rectangular wave guide.[4] In both cases experimental results have been consistent with the theoretical predictions.

In dealing with the expression in (4), no attempt was made to obtain A and B by fitting the solution to the boundary, but the form (3), in which A and B do not appear explicitly, was utilized directly.

In any particular case it is necessary to confirm the absence of an E_y component arising from the boundary conditions. Thus, since the tangential component of electric field must vanish at the aperture rim, the method will not work with a circular aperature; and in fact a rectangular opening is the only form for which the solution is appropriate.

[1] J. H. Crysdale, J. Appl. Phys. **25**, 270 (1954).
[2] G. Bekefi, J. Appl. Phys. **24**, 1123 (1953).
[3] L. Lewin, *Advanced Theory of Waveguides* (Iliffe and Sons, London, 1951), p. 121.
[4] See reference 3, p. 88.

The Electron Atmosphere in Dielectrics

M. J. MORANT
High Voltage Laboratory, Queen Mary College, London, England
(Received January 28, 1954)

IN a recent paper, Skinner, Savage, and Rutzler,[1] have considered the electron atmosphere in dielectrics bounded by plane metal electrodes. Their calculations proceed from earlier work by Mott and Gurney,[2] in which the potential $V(x)$ due to the space charge, of density $\rho(x)$, in the dielectric is calculated from the Poisson and Boltzmann equations.

$$\nabla^2 V = \frac{4\pi\rho}{K} \tag{1}$$

$$\rho = \rho_0 \exp\frac{e(V-V_0)}{kT}, \tag{2}$$

where K is the dielectric constant, and ρ_0 and $V_0=0$ are fixed by the metal-dielectric boundary ($x=0$), where

$$\rho_0 = 2e\left(\frac{2\pi mkT}{h^2}\right)^{\frac{3}{2}} \exp\frac{-e(\phi-x)}{kT}. \tag{3}$$

Mott and Gurney show that

$$V = -\frac{2kT}{e}\log_e\left\{\left(\frac{2\pi e\rho_0}{K\cdot kT}\right)^{\frac{1}{2}}x+1\right\}, \tag{4}$$

for a single metal in contact with an infinite dielectric. This was compared by von Laue[3] with the arrangement of two electrodes a finite distance apart, for which

$$V = -\frac{2kT}{e}\log_e\cos\left(\frac{2\pi e\rho_0'}{K\cdot kT}\right)^{\frac{1}{2}}x. \tag{5}$$

In this $x=0$, $V=V_0$ at a point mid-way between the electrodes, and the corresponding charge density ρ_0' has been given in terms of ρ_0 of Eq. (3) by Fowler.[4] It was also shown by Fowler that (5) approximates to (4) for sufficiently small charge densities. Skinner, Savage, and Rutzler use a generalization,

$$V = -\frac{2kT}{e}\log_e(ax+b) \tag{6}$$

of Eq. (4) to allow for different boundary conditions, but this form of solution is not permissible because of the lowering of the work function in the presence of a dielectric and the larger mean charge densities involved. This can be seen from Fig. 1 which gives a

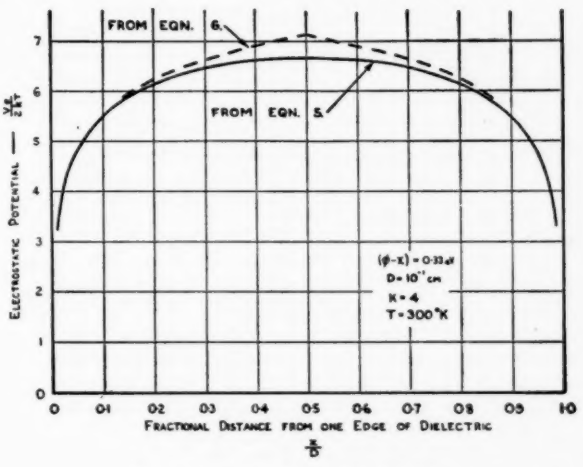


FIG. 1. Potential in a bonded dielectric.

comparison of (5) and (6) for a typical case quoted by them. The correct solution removes the anomaly of a discontinuity in dV/dx at the midpoint which is shown in many of their curves.

1. Also there seems to be some confusion by these authors in their use of electron energy level diagrams. For more than one dielectric it is necessary to distinguish between the potential energy of an electron in the dielectric and the potential as given by the solution of the electrostatic problem. The energy of an electron at the bottom of the conduction band (depth x) of an insulating crystal, and at a point of electrostatic potential V is $-(eV+x)$. Thus the boundary conditions for a pair of electrodes of the same metal at a spacing D, are

$$V(0)=0, \quad V(D)=0, \tag{7}$$

and not

$$V(0)=-x_1, \quad V(D)=-x_2. \tag{8}$$

For different metals of work functions ϕ_1 and ϕ_2 if $V(0)=0$, then

$$V(D)=\frac{1}{e}(\phi_2-\phi_1), \tag{9}$$

i.e., equal to the contact potential between the metals. This seems to have been ignored in reference 1, Fig. 3b of which shows a difference of Fermi levels of the metals in equilibrium with the dielectric. The equality of the Fermi levels of all metals in thermal and electrical equilibrium is independent of the presence of a dielectric between them, and is brought about by the contact PD. This also follows from Eq. (5) when correct boundary conditions are used.

At the contact of two insulating crystals, the equations given in reference 1 lead to a change in the level of the conduction band of

$$\phi_1-\phi_2+kT\log_e\left(\frac{K_1}{K_2}\right)$$

(not $2kT\log_e(K_1/K_2)$ if the derivation is followed through completely). This arises from the use of the wrong Eqs. (6) and (8). The complete solution for the two-electrode, two-dielectric case following Fowler, leads to a different, more complex, equation equation for V. It can be seen, however, that V should be concontinuous at the dielectric-dielectric interface so that the potential energy jump at the bottom of the conduction band is just (x_1-x_2). Using these facts it is possible to show diagrammatically, Fig. 2, the energy levels across the whole space.

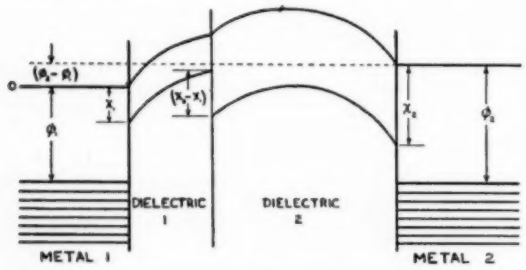


FIG. 2. Electron energy levels for two dielectrics in equilibrium with the electrodes.

2. Skinner, Savage, and Rutzler consider the equilibrium of a dielectric bonded to one electrode and with a small air gap to the second electrode. Figure 2 shows that the form of the potential variation will then be different from the doubly bonded case for all values of x if the air gap is sufficiently thick for tunnelling to be unimportant. However, the total potential between the electrodes will remain at the contact potential difference and will not approach infinity as indicated by $\rho/\rho_0\sim 0$. This affects the electrostatic contribution to adhesion more than the other errors.

3. As for frictional electrification, the measured charges may not be entirely explained by this process, because of low values, of the order of 0.1 ev, deduced for $(\phi-x)$. For such values, the charges measured should be extremely sensitive to variation of metal. The only theoretical estimates of x (those of Mott and Gurney[2]) give approximately 0.5 ev, so that $(\phi-x)$ should not be less than 2 ev.

4. A further point that may easily be overlooked in considering dielectric space charges is the effect of the Fermi level of the crystal. Mott and Gurney's calculation was for an ideal crystal (i.e., an insulator because of lack of conduction electrons), so that the distance E_F of the Fermi level below the conduction band was very much greater than the increase, within a reasonable distance from the electrode, of ev due to the space charge. Thus the naturally occurring conduction electrons could be completely neglected. Applying the same model to an amorphous insulator, E_F may be less, although still in the center of the forbidden range of energies, since the insulating property can be due to low electron mobility rather than lack of conduction electrons. This limits the potential rise, due to the space charge, to a value E_F/e so that the bulk metal and dielectric have their Fermi levels at the same height. The space charge is set up at the metal-dielectric contact to bring about this equilibrium condition. If E_F is low the whole space charge may be concentrated near the electrodes, so that calculations based on an ideal crystal may be misleading. Similar considerations have to apply to the space charge at the contact of two dielectrics.

As pointed out by Skinner, Savage, and Rutzler, electron traps which must strictly be included for an amorphous solid would modify the model considerably and lower the numerical value of E_F even more. The voltage limitation considered above would be much more pronounced in such cases.

[1] Skinner, Savage, and Rutzler, J. Appl. Phys. **24**, 438 (1953).
[2] N. F. Mott and R. W. Gurney, *Electronic Processes in Ionic Crystals* (Oxford University Press, London, 1948), second edition.
[3] M. von Laue, *Handbuch der Radiologie*, 6, (Leipzig, 1925), p. 460.
[4] R. H. Fowler, *Statistical Mechanics* (Cambridge University Press, Cambridge, 1936), second edition, p. 367.

Electrical Phenomena in Adhesion. I. Electron Atmospheres in Dielectrics

D. C. WEST
Nova Scotia Research Foundation, Halifax, Canada
(Received February 25, 1954)

THE recent article by Skinner, Savage, and Rutzler[1] under the above title appears to be a very valuable and thought-provoking study. However, there are one or two points which seem to have been overlooked.

It has been assumed implicitly that there is no formation of surface charge at the junction of two insulators, as indicated by the boundary conditions imposed by Eq. (2) of the paper. The more general boundary conditions are

$$\lim_{\epsilon\to 0}[K_1F(d-\epsilon)-K_2F(d+\epsilon)]=-4\pi\sigma, \tag{1}$$

where K_1 and K_2 are the dielectric constants of the two media which meet at $x=d$, $F(x)$ is the electric field, and σ is the surface charge density at the boundary. Since there is an electron atmosphere in both media, it would be well to consider the possibility of an accumulation of charge at the surface of discontinuity. It is known, for example, that relatively large amounts of "static" electricity can be transferred from one dielectric to another by simple contact.

An application of the same general equation to the situation at α, where the field reverses sign within one of the dielectrics, shows that

$$\lim_{\epsilon\to 0}[F(\alpha+\epsilon)-F(\alpha-\epsilon)]=4\pi\sigma/K. \tag{2}$$

Comparison with Eq. (3) of the paper shows that if there is no discontinuous charged surface at α(i.e., $\sigma=0$), the factor g must be equal to -1. The case $g=+1$, treated in the paper, can occur only if there exists a discontinuous layer of charge at α with density $\sigma=KF(\alpha+\epsilon)/2\pi$. Although mathematically correct, their results, therefore, apply to a somewhat unreasonable physical situation.

Fortunately it is quite easy to find another solution which does not require the presence of a thin sheet of charge in the center of a solid dielectric. Following the derivation given in footnote 10 on page 442 of the paper, we have from diffusion theory and Einstein's relation

$$NeF=kTdN/dx.$$

From the electrostatic field equations we also have

$$KdF/dx=4\pi Ne.$$

Elimination of N from these two equations yields

$$kTd^2F/dx^2-eFdF/dx=0, \tag{3}$$

which on integration gives

$$2kTdF/dx-eF^2=\text{constant}. \tag{4}$$

There are now three cases, depending on whether the constant of integration is (a) zero, (b) positive, or (c) negative.

(a) The choice of zero gives the solution of Mott and Gurney[2] used in Skinner, Savage, and Rutzler's paper. It is most suitable for the case of a semi-infinite dielectric in zero external field.

(b) Setting the constant equal to $4k^2T^2/e^2m^2$, the solution becomes

$$F=-(2kTm/e)\cot m(x+x_0), \tag{5.1}$$

where x_0 has the same significance as in the paper, and m is an additional parameter. The space charge N and potential energy eV become

$$N=(KkTm^2/2\pi e^2)\operatorname{cosec}^2 m(x+x_0) \tag{5.2}$$

$$eV=2kT[\log\sin m(x+x_0)-\log\sin mx_0]+eV_0. \tag{5.3}$$

It will be observed that the solution is valid only in a region $-x_0<x<-x_0+\pi/m$ and that near these limits the space charge, field and potential all tend to become infinite. This corresponds to the case of a dielectric slab of finite thickness between two sources of charge. x_0 depends on the charge density N_0 at $x=0$ as in the paper, and we may also introduce a corresponding constant x_1 connected with the charge density N_1 at the junction of the dielectric with the second metal at $x=D$. The exact relationships are

$$mx_i=\operatorname{cosec}^{-1}(e/m)\sqrt{(2\pi N_i/KkT)} \tag{6}$$

where $i=0$ or 1. The potential difference across the slab is given by

$$V_1-V_0=(kT/e)\log N_1/N_0. \tag{7}$$

The additional parameter m is connected with the dielectric thickness D by the relation

$$m(D+x_0+x_1)=\pi. \tag{8}$$

All the functions are continuous within the dielectric as required for a reasonable physical theory. They tend toward the form given in the paper as $D-\infty$.

(c) In a similar way the choice of a negative integration constant leads to a third solution, corresponding to the case of a semi-infinite dielectric subjected to an external field. At large distances this field reduces to the value $F_\infty=2kTm/e$. In this case m becomes a parameter characterizing the field strength. This solution may possibly find application to the study of dielectric absorption in charged condensers.

[1] Skinner, Savage, and Rutzler, J. Appl. Phys. **24**, 438 (1953).
[2] Mott and Gurney, *Electronic Processes in Ionic Crystals* (Oxford University Press, London, 1940), p. 172.

Electrical Phenomena in Adhesion: Further Comments on Electron Atmospheres*†

SELBY M. SKINNER, ROBERT L. SAVAGE, AND JOHN E. RUTZLER, JR.
Department of Chemistry and Chemical Engineering, Case Institute of Technology, Cleveland, Ohio
(Received May 10, 1954)

AS Dr. West says, the diffusion equation has 3 solutions;[1] these were considered in preparing the paper. In their application to the *general* single dielectric, and even more so, to two dielectrics between the metals, the "finite" solutions become excessively complicated.[2] Since the electrostatic contribution to adhesion is only appreciable with boundary charge densities of such magnitude that the infinite solution and the finite solution are essentially the same, the former was used as an approximation. Physically this means that the charge distribution characteristic of the infinite case has been chopped off at negligible charge, i.e., at a distance corresponding to an appropriate thickness of dielectric, and a similar distribution placed on the other end. This chopping off is responsible for the discontinuity in slope of the potential and charge density.

For adhesive thicknesses encountered in practice ($>10^{-4}$ cm), and electron densities ($>10^{20}$ electron cm^{-3}) at which the electrostatic contribution need be considered, the effect on the adhesive force is in error by 0.001 percent or less, and the distributions within the adhesive are affected by a similar negilgible fraction of the experimental uncertainties in adhesion measurements.[3] However, the discontinuous charge distribution at $x=\alpha$ was overlooked in determining the total charge within the dielectric. Its effect is to change the expressions for the surface charge density on the metal,[4] which should read N_0ex_0, and to cause the adhesive contribution to be N_0kT, both being irrespective of the thickness of the dielectric either before or after break. This requires obvious alterations to Eqs. (6) to (8.3) inclusive, and minor alterations in the discussion of the work of Meissner and Merrill, in the second and third paragraph of V, 2, and in section VI, 1, where the words "repulsion" and "reversal of charge" should be replaced by "enormously reduced attraction"[5] and "greatly reduced charge," respectively. In section VI, 1, the discussion on the dependence of the force on the sample thickness is inapplicable.

2. As Mr. Morant says, the factor $2kT\ln K_1/K_2$ in the potential jump between the dielectrics (at $x=d$) should be divided by two if our computation is followed through. However, his discussion (sections 2, 3, and 5) of voltages, Fermi levels, etc., does not take into account the full thermodynamic picture, and confuses Fermi levels with inner electrostatic potentials. Either the Volta potential or the inner potential can be used, but whichever one is chosen must be followed through consistently. For the Volta potential, the boundary conditions are (Fig. 1) $\mathcal{V}_1(0)=0$, $|\mathcal{V}_2(D)|=\varphi_2-\varphi_1$ (not zero), and the usual electrostatic conditions on potential and field at the dielectric-dielectric boundary. If the inner potential is used, the difference at the latter boundary is $|\chi_2-\chi_1|$ if the three strictly accurate sets of transcendental equations are used. But then it is necessary to consider the thermodynamics of the metal-dielectric boundary:

(a) With no external applied potential, the Fermi level (being the free energy per particle) equalizes throughout both metals and dielectric. The Volta potential difference in the metals raises the energy of the electron vapor outside one metal with respect to the other; the external electron densities are related by a Boltzmann factor containing this energy difference, namely:[6]

$$\left.\begin{aligned}&|e\mathcal{V}|=\varphi_2-\varphi_1;\ N_{20}/N_{10}=\exp[-(\varphi_2-\varphi_1)/kT];\\ &N_{j0}=N_e\exp[-\varphi_j/kT];\ N_e=2(2\pi mkTh^{-2})^{\frac{3}{2}}\end{aligned}\right\}. \tag{1}$$

(b) In a nonmetal, when a space charge exists, $n=n_i\times\exp[(E_F-\zeta)/kT$, where ζ is the intrinsic Fermi level, and E_F the quasi-Fermi level (imref); therefore,

$$\mathcal{V}=\zeta-kT\ln n_i+c=E_F-kT\ln n+c, \tag{2}$$

in which $\mathcal{V}$ is the electronic charge times the inner electrostatic potential, and c is the same for all substances in thermodynamic equilibrium. If there is no space charge, a rise in the Fermi level means an increased electrostatic potential. With space charge, the constant height of the Fermi level means that the concentration n of electrons toward the interior decreases just enough to make up for the increase of the inner electrostatic potential due to space charge.

Since the electron densities are different at the two boundaries, the inner electrostatic potentials will be different there. By generalization of Eqs. (3) to the atmosphere in a dielectric and use of Eq. (2) and the value $c=kT\ln N_e$, we have

$$\mathcal{V}_1(0)=E_F+(\varphi_1-\chi_1);\ \mathcal{V}_2(D)=E_F+(\varphi_2-\chi_2). \tag{3}$$

Equation (3) or its equivalent obtained by adding the same constant to both potentials is the correct boundary condition on the inner potentials, and c cannot make both potentials zero.

Equations (3) and the requirement that the potential be continuous and have a maximum in one dielectric must be satisfied. With the "finite" equations, there then are specified the potential jump at $x=d$, and the relation of the electrical

fields. With the approximate solution the maximum is no longer smooth; also it is necessary to relax one condition. A double layer or a charge sheet must be permitted at $x=d$; the latter introduces considerable error, the former negligible error, so it is chosen. Using these boundary conditions, the expressions in the paper are obtained, except that the right-hand side of (3) above appears in the constants in the potential expressions; the space dependence of the inner potentials, the fields, and charge densities, and the computed adhesive forces are those of the paper. The energy gap is $kT \ln K_1/K_2$ independent of either φ or χ. The deisred quantities, namely charge transferred and adhesive forces, are accurate to within a fraction of experimental error. The difference from the solution in the paper does not affect the adhesion to the metal, but takes account of a transfer of electrons between the dielectrics themselves; this equalizes Fermi levels in them and produces an electrostatic attraction between them if they are solids.

3. See Morant's Sec. 5: By the expressions, the maximum inner potential difference (with respect to the interface) attainable within even a 2-cm thick layer of adhesive at room temperature and minimum x_0 is 0.75 v and reasonable values are 0.1–0.7 v. The energy gap in insulators is from 2 to 7 volts. However a more fundamental point is involved: the charge density decreases as the inner potential increases, and the Fermi level remains constant; in the general nonmetal, it is known that electrons are raised thermally to conduction levels only if the Fermi level moves up. Field emission requires much higher (externally applied) fields and a decreasing rather than increasing inner potential. Commercial insulators and plastics are insulators precisely because precaution has been taken to avoid impurity levels in the intrinsic gap. Thermal contributions are not therefore a limitation on the inner electrostatic potential in insulators, though they may be in semiconductors.

4. With respect also to Morant's Sec. 5, mathematically it is possible to assume insulator resistance to be due to small charge concentration or small mobility. It is quite possible that the mobility is smaller than in semiconductors, but to attribute the small conductivity entirely or even in major part to smaller mobility raises conceptual difficulties. (1) Astronomical periods are required to reach equilibrium. (2) Even amorphous substances have molecular structure. Since polystyrene, for example, has 10^{-18} the conductivity of germanium, attributing the insulating value to poor mobility requires assuming an electron mean free path 10^{-18} that in germanium, i.e., less than 10^{-22} cm. Such a value has little physical significance; even reasonable modifications of assumed electron temperature do not correct the difficulty. Electron concentration in the conduction level, however, depends exponentially on the intrinsic energy gap; a gap of 0.7 to 2 ev would be sufficient to explain, on the basis of charge concentration, observed insulating properties. Kimball[7] has shown that in diamond the gap may be as great as 7 ev. The question can, of course, only be settled by measuring both n and b by experimental methods suitable to insulating materials.

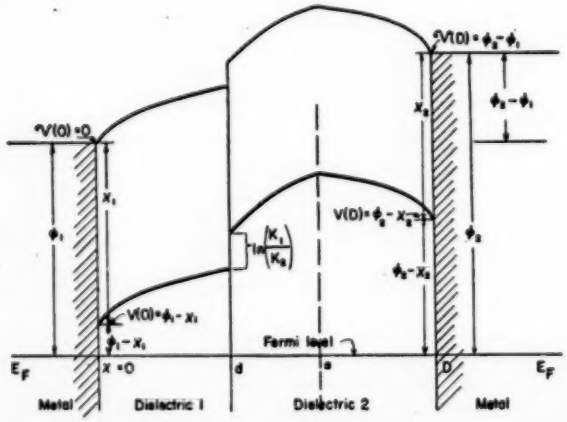


FIG. 1. Energy levels after infinite time.

5. Morant's Sec. 3 appears to refer to the legend on Fig. 5(b) (2) of (I). This is the charge density at $x=d$ not at the metal; at the metal, n goes to N_0 not 0, though at d it reaches a fraction of N_0. We agree with him that the present treatment does not explain frictional electrification, and did not make such claim.

6. The intention of the article was to point out that the transfer of electrons from a metal into a dielectric adhesive could bring about a significant electrostatic contribution to the measured adhesion. Such transfer does not depend upon the properties of the valence bonds or energy levels of the dielectric alone (which determine the intrinsic energy gap), but upon the energy gap at contact, i.e., between the Fermi level of the metal and the conduction level of the adhesive; therefore the force cannot be predicted from the measured chemical properties of the adhesive alone. It is our understanding that in at least two commercial research laboratories the subject is now being investigated in detail; the results of such investigations together with ours should yield useful data on contacts between adhesives or similar materials and metals, and help to determine the range of applicability of the present considerations.

* With the support of the Aeronautical Research Laboratory, Wright Air Development Center, U. S. Air Force.

† Since much of what follows was in manuscript when the preceding letters were received, their acknowledgment was to be incorporated therein. A delay in the second article makes it desirable to treat the points raised, directly. For lack of space much of the contents of correspondence to Dr. West and Mr. Morant is omitted and will be included elsewhere. The first article by the present authors is designated by (I).

[1] Treated by v. Laue (as mentioned above); Richardson (references in Fowler, *Statistical Mechanics*); Fowler himself; Mott and Gurney, *Electronic Processes in Ionic Crystals* (Oxford University Press, London, 1948) second edition; and F. Borgnis, Z. Physik 100, 117 (1936).

[2] S. M. Skinner, J. Appl. Phys. (to be published).

[3] This is discussed more fully in the second article on the contributiont of electron atmospheres to adhesion, Skinner, Savage, and Rutzler, which will be submitted in the near future to J. Appl. Phys.

[4] The additional surface charge density at the metal due to the discontinuous charge sheet is at most 4 percent and usually of the order of 10^{-3} or less.

[5] The electrostatic component of adhesion at room temperature between a metal surface coming up in air to one side of a thin dielectric film already adhering on its other side to a metal surface will be far less than one microdyne cm^{-2}, whereas the electrostatic component of the adhesion of the film to the first metal could be 600–1000 lb in.$^{-2}$.

[6] See R. H. Fowler and E. A. Guggenheim, *Statistical Thermodynamics* (Macmillan Company, New York, 1939), p. 486.

[7] G. E. Kimball, J. Chem. Phys. 3, 365 (1935).

Direct-Current Transients in Polymethyl Methacrylate and in Polystyrene

PAUL EHRLICH*
Diamond Ordnance Fuze Laboratories,† Washington 25, D. C.
(Received February 17, 1954)

QUANTITATIVE measurements of the dc transients in materials with dc conductivities of the order of 10^{-18} mho/cm and less have not been reported previously. This communication reports on such measurements made with a commercial Vibrating Reed Electrometer[1] on two widely used high polymer insulating materials. The procedure used was that of plotting on an automatic recorder the charge accumulated on the internal capacitor of the instrument. The voltage across this capacitor was 1 volt maximum which was only a small fraction of the battery voltage. The effect of the measuring instrument on the normal current flow through the specimen was therefore negligible. The samples were molded disks $\frac{1}{16}$ in. thick with an effective electrode diameter of approximately 3 in. Data are presented for two different brands of commercial polymethyl methacrylate, a sample of polystyrene vacuum-polymerized at the National Bureau of Standards, as well as a commercial sample of the same polymer. Figure 1 is an illustration of the circuit used.

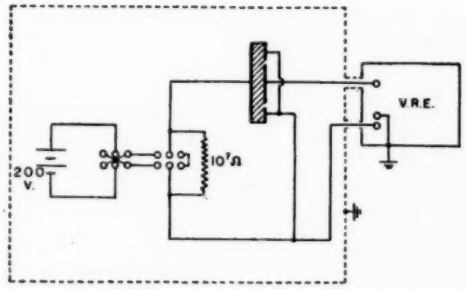


FIG. 1. Circuit used in conjunction with Vibrating Reed Electrometer (V.R.E.) for the measurement of very low dc conductivities.

Figures 2 and 3 present data of charging and of discharge currents in polymethyl methacrylate and in polystyrene as a function of time. It is seen that in all cases measurable absorption currents continue to persist over the entire time scale and that, even at the longest times, it is impossible to report steady-state conductivities. In one of the polystyrene samples the steady-state conductivity must be only a small fraction of the smallest transient conductivity observed, as can be seen from the identity of charging and discharge curves over the entire time scale.[2] In the other two cases, the charging curves bend upward and away from the discharge curves at the upper end of the time scale, indicating that the steady-state conductivities have become a substantial fraction of the total conductivities. There is a significant difference in the data for the two samples of polystyrene; data obtained for a different commercial brand of polymethyl methacrylate, however, gave, within experimental error, identical results with the ones presented, except that the charging curves did not bend away from the discharge curves at long times, indicating the absence of any substantial steady-state conductivity.

It is of interest to examine the connection between the measurements reported and the ac loss factor. It is well known that a relation exists between the frequency-dependent dielectric constant $\epsilon'(\omega)$ the loss factor $\epsilon''(\omega)$ and the dc transient current, $i(t)$.[3] Von Schweidler[4] has calculated apparent capacitances and resistances for a decay function of the form $i(t)=At^{-n}$ and $\epsilon'(\omega)$,

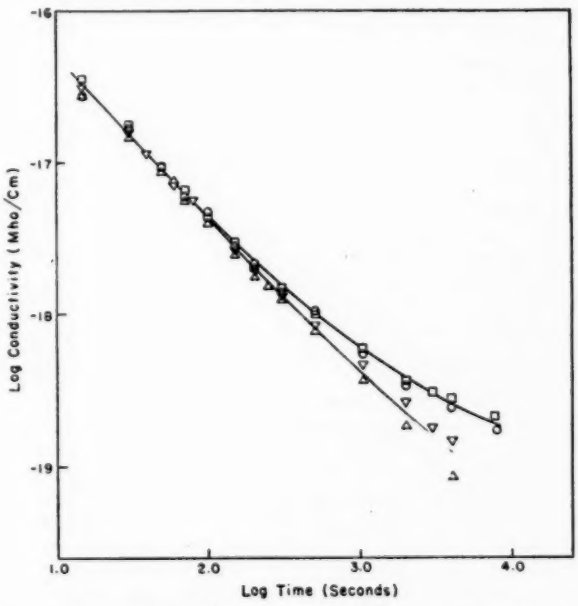


FIG. 2. Logarithm of the dc conductivity of a sample of polymethyl methacrylate plotted as a function of the logarithm of the time after application (○,□) and removal (△,▽) of the potential. The second symbol in each case refers to a run made under reversal of the battery terminals.

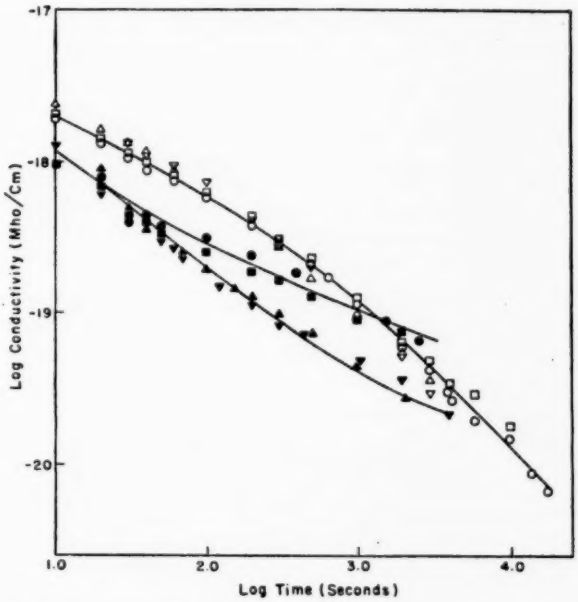


FIG. 3. Logarithm of the dc conductivity of a commercial sample of polystyrene plotted as a function of the logarithm of the time after application (○,□) and removal (△,▽) of the potential. The second symbol in each case refers to a run made under reversal of the battery terminals. The filled in symbols refer to a sample of vacuum-polymerized polystyrene.

as well as $\epsilon''(\omega)$, can be immediately obtained from this development with the following result for the loss factor:[5]

$$\epsilon''(\omega)=1/C_a[G_0/\omega+\beta C_0\omega^{n-1}\Gamma(1-n)\cos(n\pi/2)], \quad (1)$$

where C_a is the capacitance of a sample of air which has replaced the dielectric, C_0 is the capacitance of the dielectric at very high frequencies, G_0 is the steady-state dc conductivity, and A, β, and n are constants. It is readily seen that the loss factor may increase, decrease, or remain constant with frequency, depending on the exponent n of the decay function. By a suitable variation of the exponent through the time scale, the loss factor may be made to assume any of the shapes ordinarily observed. Hamon has shown[5] that, subject to certain requirements almost invariably satisfied, Von Schweidler's method can be made the basis of an approximation which is valid even where the exponent is not constant over the entire time scale. Hamon obtains

$$\epsilon''\simeq i(0.1/f)/2\pi f C_a V, \quad (2)$$

which may be written as

$$\epsilon''\simeq G(0.1/f)/2\pi f\epsilon_0, \quad (3)$$

where $i(0.1/f)$ and $G(0.1/f)$ are transient current and transient conductivity, respectively, at a time equal to one tenth of the reciprocal cps frequency, V is the applied voltage and ϵ_0 is the permittivity of free space (8.854×10^{-14} farads/cm).

Table I presents values of the loss factor for polymethyl methacrylate and polystyrene calculated from the transient measurements by Eq. (3), and values at higher frequencies obtained

TABLE I. Loss factor of polymethyl methacrylate and polystyrene.

Polymethyl methacrylate		Polystyrene Vacuum-polymerized		Polystyrene commercial sample	
Frequency (cps)	Loss factor	Frequency (cps)	Loss factor	Frequency (cps)	Loss factor
10^{-3}	8.1×10^{-3}	10^{-3}	3.2×10^{-4}	10^{-3}	$11.\times10^{-4}$
6.7×10^{-3}	9.3×10^{-3}	10^{-2}	2.0×10^{-4}	10^{-2}	3.6×10^{-4}
10^{2}	2.1×10^{-1}	10^{2}	2.5×10^{-4}	10^{2}	2.8×10^{-4}
10^{3}	1.4×10^{-1}	10^{3}	2.5×10^{-4}	10^{3}	2.8×10^{-4}

directly on a Schering bridge. Barring a sharp dispersion region between 10^2 and 10^{-2} cps, polystyrene would therefore appear to undergo no very decided change in loss factor at room temperature over most of the interval considered. Polymethyl methacrylate, on the other hand, has to undergo roughly a 20-fold drop in loss factor in a little over four frequency decades. Judging from the behavior of this polymer at other temperatures and frequencies,[6] this value would seem quite high; nevertheless, polymethyl methacrylate has been shown to have a 100 cps loss factor maximum at about room temperature[6] and a marked drop in loss factor on the low-frequency side of this maximum should, therefore, be expected.

The author gratefully acknowledges helpful advice given to him by Dr. A. H. Scott and Mr. N. Doctor in setting up the measuring circuit. He is also indebted to Dr. L. A. Wall for providing the vacuum-polymerized polystyrene.

* Present Address: Department of Chemistry, Harvard University, Cambridge 38, Massachusetts.
† Formerly a part of the U. S. National Bureau of Standards.
[1] Palevsky, Swank, and Grenchik, Rev. Sci. Instr. **18**, 298 (1947).
[2] A dc conductivity of 10^{-20} mho/cm for a very pure polystyrene 45 minutes after application of the potential has just been reported in another investigation. Warner, Muller, Nordlin, J. Appl. Phys. **25**, 131 (1954).
[3] K. S. Cole and R. H. Cole, J. Chem. Phys. **10**, 98 (1942).
[4] E. Von Schweidler, Ann. phys. **24**, 711 (1907).
[5] B. V. Hamon, Proc. Inst. Elec. Engrs. **99**, Part IV, 151 (1952).
[6] D. J. Mead and R. M. Fuoss, J. Am. Chem. Soc. **64**, 2389 (1942).

Optimal Loading of a Thermoelectric Generator

JAMES F. PERKINS
Convair, Fort Worth, Texas
(Received March 3, 1954)

TELKES[1] has derived an equation for the over-all efficiency of a thermoelectric generator, for which it was assumed that the maximum over-all efficiency would be obtained by matching the load resistance R_e to the generator internal resistance R. Papet[2] has pointed out that, while for a given generator the maximum power output can be achieved by matching impedances, the optimal operating condition is one corresponding to a maximum over-all efficiency and that this can be obtained by using a load resistance, R_e, somewhat greater than the internal resistance R of the tenerator. Following Papet's suggestion, Greeff[3] has calculated efficiencies for various loading conditions. Some of the calculated efficiencies are in error by more than a factor of two, however, as a result of neglecting the Peltier effect. For instance, if one analyzes generator II according to the reasoning applied to cases III, IV, and V, one concludes that, since generator II is only 1/25 as large as generator I, it will absorb only 1/25 as much heat at the hot junction. Since the outputs of I and II are the same as we have: Efficiency of II = 25 × Efficiency of I = 25 × 2.34 percent = 58.8 percent. By comparison, the actual efficiency of II was calculated by use of Telkes' formula to be 20.9 percent.

If one generalizes Telkes' equation to include the value of the load resistance, one obtains

$$\text{Efficiency} = \frac{1}{A\left(\frac{R+R_e}{R_e}\right)+B(R+R_e)^2/(RR_e)},$$

where

$$A = 1/(\text{Carnot efficiency}) = T_H/(T_H-T_C).$$
$$B = \frac{4.92\times10^{-8}(T_H+T_C)}{e^2(T_H-T_C)}.$$

The condition for maximum efficiency, which is easily obtained by differentiating with respect to R_e/R, is

$$(R_e/R)_{\text{optimal}} = (A+B)^{\frac{1}{2}}/B^{\frac{1}{2}}.$$

In the limiting case of $A/B \to 0$, i.e., the case of vanishing efficiency, the Peltier term does not contribute to the heat transfer. Hence, maximum efficiency corresponds to maximum output, and thus is obtained by impedance matching, $R_e = R$. In Greeff's Case I, for example, we find $A/B = 1.546/10.3 = 0.15$, and $(R_e/R)_{\text{optimal}} = 1.075$.

The limiting case $A/B \to \infty$ corresponds to a generator whose thermal partial efficiency, as defined by Papet, is independent of R_e. Over-all efficiency is thus proportional to electrical efficiency $\eta_e = R_e/(R+R_e)$ and is maximized by choosing R_e much larger than R, i.e., $(R_e/R)_{\text{optimal}} \to \infty$.

In Greeff's case II we find $A/B = 1.546/0.412 = 3.74$, $R_{e\ \text{optimal}} = 2.18R = 4.36$ ohms, $\text{Efficiency}_{\text{optimal}} = 24.0$ percent as compared to $\text{Efficiency}_{R_e=R} = 20.9$ percent.

At still higher efficiencies the improvement in efficiency obtained by proper loading becomes more significant. For a generator with $e = 1000\mu V/°C$, operated at the temperature difference of 500°C considered by Greeff, we have $\text{Efficiency}_{R_e=R} = 28.6$ percent, $\text{Efficiency}_{\text{optimal}} = 39.0$ percent, $(R_e/R)_{\text{optimal}} = 4.0$.

[1] M. Telkes, J. Appl. Phys. **18**, 1116 (1947).
[2] R. M. Papet, J. Appl. Phys. **19**, 1180 (1948).
[3] M. B. Greeff, J. Appl. Phys. **21**, 943 (1950).

Power Output of Thermoelectric Generators

MARIA TELKES
Research Division, College of Engineering, New York University, New York, N. Y.
(Received May 25, 1954)

THERMOELECTRIC generators—theoretically—may offer numerous advantages, when compared with conventional electric power generators. There are no moving parts, nothing to wear out, if thermal deterioration at the hot junction could be avoided. The efficiency of the modern power generating stations is around 30 percent. The writer's calculations[1] indicated that such efficiencies may be approached, if thermocouple materials are available giving a thermoelectric power of 1000 microvolt/°C (one millivolt per °C), operating at a temperature difference of 500°C, provided that the thermocouple materials have "normal" Wiedemann-Franz-Lorenz relation. This relation is expressed by

$$k/\sigma = k\cdot\rho = 2.45\ 10^{-8}T,$$

where k = specific heat conductivity, ρ = resistivity, σ = specific electric conductivity, and T = absolute temperature. When the thermocouple is operated in a temperature range with T_h hot junction and T_c cold junction temperature, the corresponding k, σ, or ρ values should be used.

The writer has constructed numerous experimental thermoelectric generators, operated at various temperature ranges. The results have been summarized[1] and another article on "Solar Thermoelectric Generators" will appear in the near future.[2] The highest efficiency obtained by the writer was around seven percent, using thermoelectric materials of 300 to 400 microvolt/°C with moderate deviations from the normal Wiedemann-Franz-Lorenz relation. In addition many materials have been used, some of them with thermoelectric power values in excess of 1000 microvolt/°C (but unfavorable $k\cdot\rho$ values), but it was impossible to obtain higher efficiencies, except the values predicted by the theoretical calculation.

Papet[3] and Greeff[4] published additional calculations indicating that the efficiency of thermoelectric generators could be increased by structural changes and by variations in the external load resistance, R_e as compared with the internal resistance R_i. Anyone who has constructed thermoelectric generators would naturally wish to obtain the maximum power output and for this reason would change the R_e/R_i ratio in the widest possible range. The writer has naturally done this, and found that the maximum power output could be obtained when R_e/R_i was 1, or between 1.05 and 0.95.

Perkins,[5] in a recent letter to the editor, calculated that the R_e/R_i ratio change may increase the theoretically calculated efficiency of thermoelectric generators, provided that thermo-

electric materials are available with 1000 microvolt/°C thermoelectric power, and normal Wiedemann-Franz-Lorenz relation. Operating at a temperature difference of 500°C, the calculation indicated that the efficiency could be increased from some 29 percent to 39 percent, in the above case, by using an R_e/R_i ratio of four. In the writer's original article[1] it was stated that the calculations should be revised, for high thermoelectric power values, because the Thomson effect was neglected (but not the Peltier effect).

Unfortunately all materials with high thermoelectric power values (500 to 1000 microvolt/°C) deviate considerably from the normal Wiedemann-Franz-Lorenz relation and it appears that an efficiency greater than seven percent has not been attained. Calculations alone will not increase the efficiency; it is essential that better thermoelectric materials should be developed, which have the desirable 1000 microvolt/°C, at a normal Wiedemann-Franz-Lorenz relation. For practical purposes operating with $T_h=800°K$, $T_c=300°K$, $k\cdot\rho=13\,10^{-6}$ (as an average for the operating temperature range). If the specific heat conductivity k were as low as that of compact insulators $k=0.01$ (in watt·cm^{-1} deg^{-1}), the corresponding ρ should not be higher than 0.0013 ohm·cm as an average in the operating temperature range. Materials of this type have not been found as yet.

[1] M. Telkes, J. Appl. Phys. **18**, 1116 (1947).
[2] M. Telkes, J. Appl. Phys. **25**, 765 (1954).
[3] R. M. Papet, J. Appl. Phys. **19**, 1180 (1948).
[4] M. B. Greeff, J. Appl. Phys. **21**, 943 (1950).
[5] J. F. Perkins, J. Appl. Phys. **25**, 1058 (1954).

Thermal Diffusivity of Metals at High Temperatures

D. ROSENTHAL AND N. E. FRIEDMANN
Department of Engineering, University of California, Los Angeles, California
(Received February 27, 1954)

UNDER the above title an experimental procedure based on the modified Angstrom[1] method has been described recently.[2] It has been known since the inception of the original method by Angstrom (1861) that an application of a periodically varying temperature at one end of a long bar offers the following advantages:

(1) The temperature of the source need not be known.
(2) Radiation and convection losses need not be determined.
(3) The value of thermal diffusivity can be obtained within a temperature range sufficiently narrow, so that it becomes practically a value at a fixed temperature.

The modified method appears to possess an additional advantage of reducing the time of measurements to a single period, once the desired (mean) temperature has been attained in the whole specimen.

What appears to be less widely known is that the Angstrom method is not the only one possessing the above advantages. The method based on the use of a moving heat source[3,4] offers the same advantages and it has the added advantage of covering wide ranges of temperature in one single run. Thus, even the waiting period for the attainment of a mean temperature is eliminated.

The method is based on the establishment of a quasi-stationary state in a long bar swept by a moving heat source. During the sweep each point of the bar is subjected to a temperature rise on the approach of the source and temperature fall when the source moves away. It can be shown[4] that if a plot is made of the log temperature variation with time at a given point sufficiently removed from both ends of the bar, the slopes β and β' corresponding to the same temperature on the rising and falling portions of the plot are related to the thermal diffusivity α at this temperature as follows:

$$\alpha=v^2/(\beta+\beta'). \quad (1)$$

Here v is the velocity of the heat source.

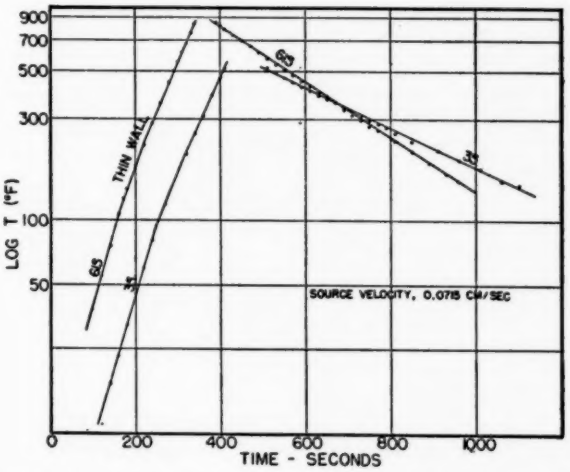


FIG. 1. Plot of time-log temperature data for 61S and 3S aluminum alloys.

As an example, two typical plots of time-log temperature data are shown in Fig. 1. These data were taken from time-temperature records using equipment similar to that described by Sidles and Danielson.[2]

Table I contains some data obtained by this method along with data secured by other methods or from other sources and included for the sake of comparison. With the present experimental technique a reproducibility of 2 percent or better appears possible in all cases in which thermal diffusivity exceeds 0.2 cm²/sec.

In spite of the advantages of both methods it is only fair to caution against undue emphasis on direct determinations of thermal diffusivity. For one thing the knowledge of thermal diffusivity alone is not sufficient to solve transient heat problems in which thermal properties vary with temperature. The variation of heat conductivity with temperature also must be given.[5,6] For another thing, accurate methods for the determination of heat conductivity and heat capacity at various temperatures are available.[1,7,8] Hence, thermal diffusivity for each temperature can be computed and the degree of accuracy appears to be at least equal, if not actually superior, to that afforded by direct methods.

The use of direct methods is therefore a matter of expediency rather than necessity. This expediency, however, is greatly

TABLE I. Thermal diffusivity of some alloys.

Alloy	Temperature Range (°C)	Thermal diffusivity (cm²/sec) Moving heat source: Value	Number of tests	Standard deviation	Other sources
Yellow brass 60/40	100–400	0.346 (1+0.000715T)	4	0.010	0.365 (1+0.001T)[a]
Deoxidized Copper	65.5–288	0.980	2	0.009	0.99[b] at 65.5°C
Aluminum Alloy 61S Nominal Composition Mg-1.0, Cu-0.25 Si-0.6, Cr-0.25	65.5–483	0.673 (1+0.0000107T)	31[c]	0.010	0.662[d] at 25°C
Aluminum Alloy 3S Nominal Composition Mn 1.2	121–538	0.643 (1+0.0000523T)	7	0.014	0.605[d] at 25°C

[a] See reference 4, 18–200°C
[b] Computed from electrical resistivity at 65.5°C using Franz-Wiedemann relation [C. S. Smith and Palmer, Trans. Am. Inst. Mining Met. Engrs. **117**, 225 (1935)].
[c] Four source velocities and two specimen thicknesses.
[d] *Metals Handbook* (The American Society for Metals, Cleveland, 1948).

justified by the present scarcity of complete sets of data on density and heat capacity over wide ranges of temperature for even the most common alloys.*

* It might be mentioned that in certain cases it is possible to evaluate the dependence of thermal conductivity and heat capacity on temperature from the known temperature variation of another physical property, e.g., electrical resistivity. Such cases are of course limited.

[1] H. S. Carslaw and J. C. Jaeger, *Conduction of Heat in Solids* (Clarendon Press, Oxford, 1948).
[2] P. H. Sidles and G. C. Danielson, J. Appl. Phys. **25**, 58–66 (1954).
[3] D. Rosenthal, Trans. Am. Soc. Mech. Engrs., **68**, 849–866 (1946).
[4] D. Rosenthal and A. Ambrosio, Trans. Am. Soc. Mech. Engrs. **73**, 971–974 (1951).
[5] M. Jacob, *Heat Transfer* (John Wiley and Sons, Inc., New York, 1949) p. 192.
[6] R. Plunkett, Trans. Am. Soc. Mech. Engrs., **73**, 605–608 (1951).
[7] G. J. Aitchison, J. Sci. Instr. **24**, 200–202 (1947).
[8] R. B. Rice, ASTM Bull. No. 157, 72–75 (1949).

Multiple Beam Sharpening of Photoelastic Fringes

DANIEL POST
Naval Research Laboratory, Washington, D. C.
(Received March 12, 1954)

CONSIDERABLE attention has recently been given to mechanical techniques for increasing the precision of photoelastic measurements. Great strides have been made possible in this direction by improved techniques in model preparation, calibration, and fringe pattern analysis, and by compensator and photometric determination of fringe orders. The present contribution deals with a technique for transforming the shape of photoelastic fringes to make them more amenable to precise measurements.

When partial mirrors of high reflectivity are placed on each side of a model, Fig. 1, part of the light traverses the model many times before emerging. For normal incidence these rays emerge from the same point, and the combined intensity varies with model birefringence according to the curve given in Fig. 2(a). The shape of the two beam interference fringes of conventional photoelastic isochromatic patterns is provided in Fig. 2(b) for comparison.

These charts were obtained by replacing the model with a Babinet-Soleil Compensator in a standard circular polariscope, and imaging the compensator aperture upon the grid of a photomultiplier tube. The compensator was driven at constant speed and the phototube output was amplified and recorded on a constant speed chart. The sharpened pattern was obtained by interposing mirrors of 85 percent reflectance (3 percent transmittance, 12 percent absorption) on each side of the compensator.

Similar charts show that errors in the optical system, such as imperfect circular polarization and extraneous birefringence

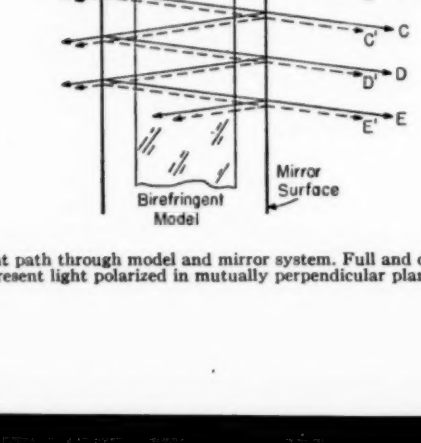


FIG. 1. Light path through model and mirror system. Full and dashed lines represent light polarized in mutually perpendicular planes.

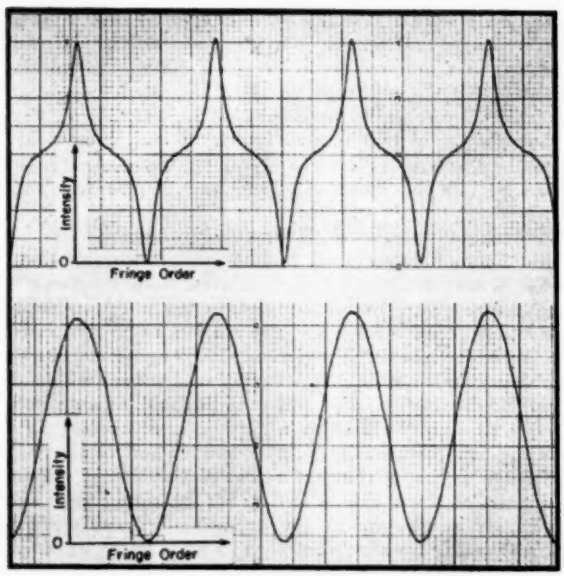


FIG. 2. Intensity curves for (a) sharpened photoelastic fringes (top) and (b) ordinary photoelastic fringes (bottom).

outside the mirror surfaces, have no effect upon the positions of the sharpened peaks. These errors reduce the amplitude of the peaks and cause minor asymmetry in fringe shape.

Rigorous mathematical analysis of this problem appears to be very difficult. However, a physical argument leads to the following intensity-retardation relationship:

$$I=\frac{I_0T^2}{2}\left[\frac{1}{1-R^2}+\frac{(1-R^2)\cos 2\pi N}{1+R^4-2R^2\cos 4\pi N}\right], \qquad (1)$$

where I is fringe intensity, I_0 is incident light intensity, T and R are fractions of light intensity transmitted and reflected at each mirror surface, and N is relative retardation, in wavelengths, caused by a single passage of light through the model. This is the exact solution for the summation of transmitted rays, assuming wave coherence for each pair of rays, $A-A'$, $B-B'$, $\cdots$, arising

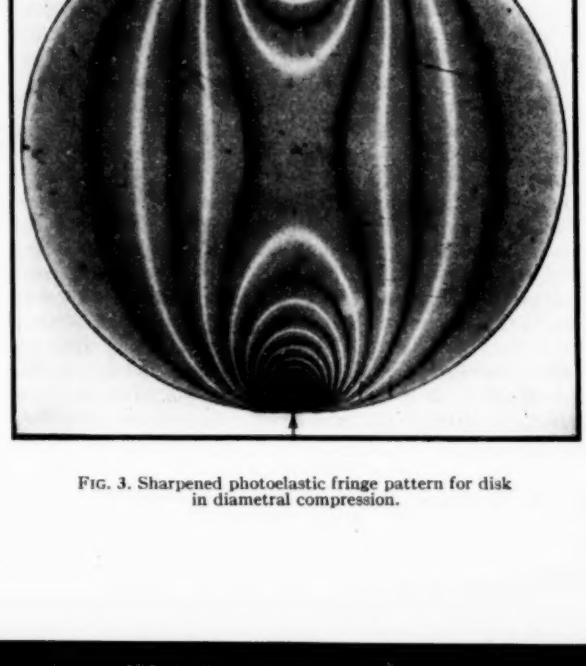


FIG. 3. Sharpened photoelastic fringe pattern for disk in diametral compression.

out of the birefringence, and noncoherence between other combinations of rays. The large distance between mirrors (in wavelengths), coupled with the finite band width of any monochromatic source, produces a randomization of phases having the essential properties and effects of noncoherent rays.

The plot of Eq. 1 yields the same shape curve as given in Fig. 2(a), but with somewhat sharper peaks. It should be remembered, however, that all imperfections in the experiment, such as finite size of source, finite spectral band, non-normal incidence, etc., produce some broadening of the sharpened fringes.

Figure 3 shows a typical sharpened photoelastic fringe pattern—a circular disk in diametral compression. The distinct advantage over conventional patterns is apparent. Fringe positions can be established with far greater precision for these multiple-beam fringes than for the corresponding system of two-beam fringes.

This technique should be most valuable for those problems where low fringe orders or low fringe gradients are important. A complete fringe distribution plot becomes possible even for these cases. Several sharpened fringe parameters can be observed and measured by utilizing monochromatic sources of different wavelengths (such as the strong lines of the mercury arc). Further, the method of varying the magnitude of applied load for the purpose of displacing fringes to given regions becomes practical with the sharpened system.

A Torsional Pendulum of Low Thermal Inertia*

C. Ang and C. Wert
University of Illinois, Urbana, Illinois
(Received March 23, 1954)

THE application of anelastic methods to problems in metallurgy and the physics of solids has been in some instances of considerable value.[1,2] Many of these applications have used, to measure the damping of the specimen, a simple torsional pendulum not much more complicated than the pendulum commonly used to measure rigidity modulus in physics laboratory courses.[3,4] Because of its simplicity, modifications in its use can easily be made to fit special needs of various problems. This account will describe a pendulum which was constructed for the study of damping in certain Au-Ni alloys.[5]

In alloys of gold and nickel there is an interesting relaxation effect which is observed when the alloys are quenched from high temperatures, but this effect disappears fairly rapidly at temperatures which one must use to measure the relaxation. Hence, to study this effect properly, an apparatus with rather low thermal inertia is required. The pendulums commonly used in the past were unsatisfactory in this respect since heat transfer from the external furnace to the specimen was too slow.[3,4] By the time a quenched specimen was mounted in the apparatus and heated to the proper temperature (some 300°C), the relaxation effect was nearly gone. Hence, some modification of the existing apparatus was required to eliminate this thermal sluggishness.

A pendulum with low thermal inertia designed to measure such transient peaks is shown semischematically in Fig. 1. The specimen (in this case a 0.030 in. diameter wire) forms the suspension of a counterbalanced inverted torsional pendulum (counterbalanced so that the longitudinal stress on the specimen is small). The wire is heated by the Joule heat of a current supplied by a variable low-voltage ac source. Electrical contact is made directly at the bottom (fixed) end of the sample and through a pin dipping into a cup of gallium at the top (free) end. The heating rate can easily be made to be several hundred degrees C per minute; at any temperature the specimen can be maintained rather constant over long periods of time.

The temperature of the specimen is measured with a thermocouple spot welded directly to the wire. The thermocouple, which

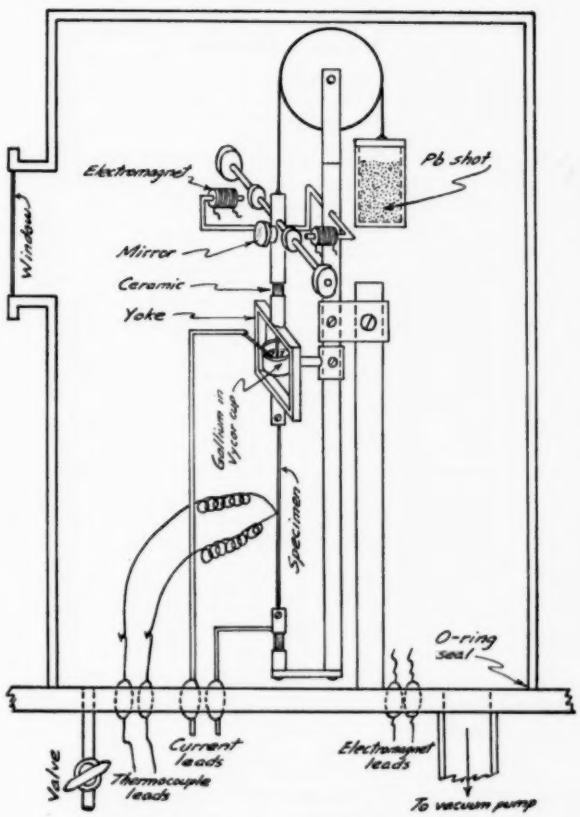


Fig. 1. A schematic drawing of the principal features of the low-thermal-inertia pendulum.

is made of wire of 0.002 in. to 0.010 in. diameter, is made to have many long, loose spirals and loops; it then adds no appreciable damping. The most serious defect of the apparatus lies in the nonuniformity of temperature along the length of the wire. The temperature falls off slowly at first from the middle to the two ends, then rapidly toward the grips which act as rather large heat sinks. For a typical 10-cm specimen nominally at 350°C, a series of 0.002 in. Cu-constantan thermocouples showed the central 7 cm to be within 10°C of 350°C, and the last 1½ cm at each end to have a steep gradient down to the grips. By using a 0.010-in. thermocouple at the center to act as a slight heat sink at that point, the central 7-cm section was made uniform to 5°C at 350°C.

This nonuniformity of temperature is much greater than would be tolerated in an apparatus using an external heating element, but does not actually seem very detrimental. Well-known relaxation peaks in Ta were observed to be shifted slightly to higher temperatures (a few degrees) and slightly broadened. These effects could be reduced either by using longer specimens or by heating the grips in some fashion.

A good vacuum can be maintained inside the metal bell to reduce oxidation at high temperatures. With the bell removed, the entire apparatus is easily accessible. Ferromagnetic specimens must be examined with caution, since large damping of magnetic origin may be observed with high ac currents needed for heating (up to 20 amperes). This effect would apparently be eliminated by using a dc heating current.

* This work was sponsored in part by the Office of Ordnance Research, U. S. Army.
[1] C. Zener, *Anelasticity of Metals* (University of Chicago Press, Chicago, 1948).
[2] A. Nowick, *Progress in Metal Physics* (Academic Press, Inc., New York, 1953), Vol. 4, p. 1.
[3] T. S. Kê, Phys. Rev. 71, 533 (1947).
[4] C. Wert, *Modern Research Techniques in Physical Metallurgy* (Am. Soc. Metals, Cleveland, Ohio, 1953), p. 225.
[5] C. Ang and C. Wert (to be published).

Maxwell's Demon Must Remain Incommunicado

MARCEL J. E. GOLAY
Signal Corps Engineering Laboratories, Fort Monmouth, New Jersey
(Received April 2, 1954)

SZILARD[1] has indicated that Maxwell's Demon cannot function because he lacks the information required for the operation of his trap door. This concept has been further explored by Brillouin[2] who has shown that the increase of entropy incidental to the operation of a sensing system for gathering the information needed about molecule movements exceeds the negentropy thus acquired for the Demon. These studies thus dispose effectively of any speculation that the entropy of information theory is different from that of thermodynamics, except for the factor $-klg_e2$.

The purpose of this letter is to point out that even if the molecule sensing system were to operate at the expense of no entropy increase, the very process of communicating this information to the Demon would require too great an entropy increase.

This can be shown very simply when the communication system consists of an ordinary transmission channel with negligible losses. The optimum information rate of such a channel is given by the now classical expression, $Wlg[1+(P/kTW)]$, where W is the band width utilized, P is the signal power, and kTW is the noise power. For small powers, this formula indicates that the information entropy S_i communicable for an expenditure of energy E, is related to the latter by the expression

$$S_i \leqslant (E/kTlg_e2).$$

The thermodynamic negentropy communicable to the Demon is therefore limited to

$$-S_{th} \leqslant (E/T),$$

i.e., is smaller than the entropy increase which occurs when the electrical energy E utilized by the communication system is dissipated into heat at the temperature T.

It would be equally futile to attempt to communicate with the Demon by means of, say, a high-energy photon dispatched at one of many prearranged epochs (pulse time modulation). The rate of blackbody emission of photons in the frequency range $\Delta\nu$ from an area of the order of a half wavelength square is given by $\Delta\nu/e^{h\nu/kT}-1$. The precision with which the time of arrival of a photon within the prearranged frequency range $\Delta\nu$ can be determined is of the order of $1/\Delta\nu$, and the natural logarithm of the number of small time intervals of this order into which the large time intervals between photon transmissions can be divided should be of the order of $h\nu/kT$, in order to transmit information with the same energy economy as with a conventional communication channel. This indicates that the rate at which photons should be transmitted $\Delta\nu e^{-h\nu/kT}$ is of the same order of magnitude as the blackbody emission from the smallest area on which photons can be concentrated.

We must conclude, therefore, that the very postulation of a communication system, either with a stupid Demon who obeys orders only, or within a clever Demon between his molecule sensing nerves and his motor nerves, destroys his reason for being.

[1] L. Szilard, Z. Physik **53**, 840–856 (1929).
[2] L. Brillouin, J. Appl. Phys. **22**, 334 (1951).

The Dynamic Mechanical and Electrical Behavior of Polymethyl Methacrylate

DAVID TELFAIR
Earlham College, Richmond, Indiana
(Received January 28, 1954)

DURING the past ten years a considerable interest has developed in the dynamic mechanical behavior of polymers. Studies in this field are not only of practical value in the rubber and plastics industries[1,2] but are also of theoretical interest because they lead to a better understanding of the molecular processes which underlie the response of polymers to mechanical and electrical stress.[3,4]

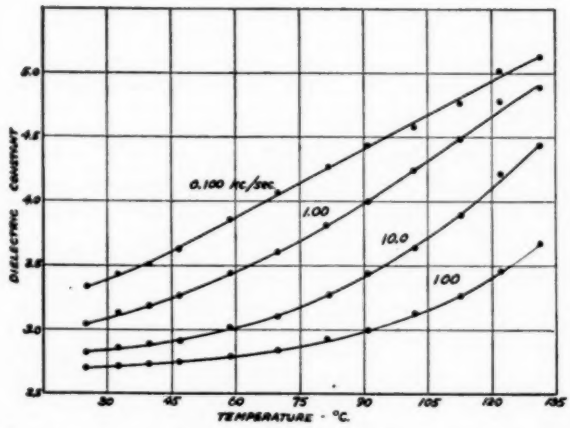


FIG. 1. Dielectric constant of polymethyl methacrylate *vs* temperature.

We have studied the effect of temperature on the mechanical damping capacity and modulus of elasticity in shear for polymethyl methacrylate by means of the torsional pendulum method described by Nielsen.[5] Dielectric constant and power factor measurements within the temperature and frequency ranges indicated in Figs. 1 and 2 have also been made.

Samples of the unplasticized polymer from two independent sources* were studied. Figure 3 illustrates the dependence of dynamic shear modulus and of mechanical damping capacity with temperature. Frequency for the torsional vibrations was held nearly constant (approximately 0.20 cycle per second).

The dielectric data shown in Figs. 1 and 2 were obtained with a General Radio type 716-C capacitance bridge. A two-terminal dielectric sample holder was used, similar to the General Radio type 1690-A holder. Mineral oil from a thermostatically controlled bath was circulated through the grounded terminal and walls of the holder, in order to control temperature. The data shown in Figs. 1 and 2 are for a predried specimen of "heat-resistant" Lucite (supplied by du Pont), but are also characteristic of the corresponding grade of Plexiglas (supplied by Rohm and Haas).

A few years ago Alexandrov and Lazurkin[6] published data on the response of polymethyl methacrylate to dynamic stressing ar various temperatures and frequencies. Fitzgerald and Ferry[7] have found that an activation energy calculated from electrical data for polyvinyl chloride is approximately the same as that calculated from mechanical damping data. An activation energy for polymethyl methacrylate estimated from the data of Alex-

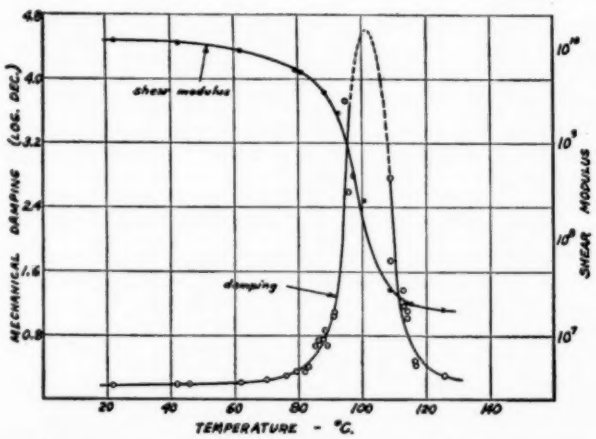


FIG. 2. Dissipation factor (tanδ) of polymethyl methacrylate *vs* temperature.

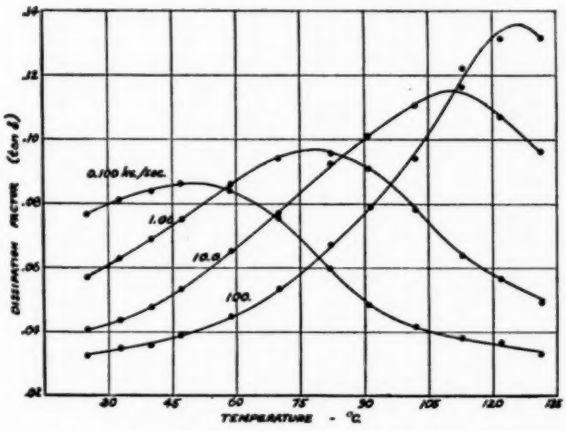


Fig. 3. Dynamic shear modulus and mechanical damping of polymethyl methacrylate *vs* temperature.

androv and Lazurkin is not in satisfactory agreement with the value calculated from the electrical data indicated in Figs. 1 and 2.

McLoughlin and Tobolsky[8] have made a study of the viscoelastic properties of polymethyl methacrylate by means of stress relaxation methods. Their data clearly exhibit the relatively strong dependence of the dynamic mechanical properties on moisture content. Some preliminary work in this laboratory indicates that the dielectric behavior is also sensitive to moisture. The results of further investigations will be reported later.

* The Rohm and Haas Company, and the E. I. du Pont de Nemours & Company, Inc.

[1] L. E. Nielsen, J. Am. Chem. Soc. **75**, 1435 (1953).
[2] W. J. Lyons and I. B. Prettyman, J. Appl. Phys. **19**, 473 (1948).
[3] Ferry, Grandine, and Fitzgerald, J. Appl. Phys. **24**, 911 (1953).
[4] Fitzgerald, Grandine, and Ferry, J. Appl. Phys. **24**, 650 (1953).
[5] L. E. Nielsen, Am. Soc. Testing Materials Bull. No. 165, 48 (1950).
[6] A. P. Alexandrov and J. S. Lazurkin, Acta Physicochim. (U. S. S. R.) **12**, 647 (1940).
[7] E. T. Fitzgerald and J. D. Ferry, J. Colloid Sci. **8**, 1 (1953).
[8] J. R. McLoughlin and A. V. Tobolsky, J. Colloid Sci. **7**, 555–568 (1952).

Hardening of Pd by Cathodic Hydrogen

TAKESI SUGENO AND MASAMICHI KOWAKA
The Institute of Industrial and Scientific Research, Osaka University, Sakai-shi, Osaka, Japan
(Received April 2, 1954)

WHEN palladium absorbs hydrogen, the two phases, α and β, are known to form depending upon the concentration of hydrogen, and they are both of f.c.c. lattices but differing in lattice parameters and exist as segregated phases.[1] It is also known that Pd is hardened by the hydrogen absorption. We have measured the hardness of the Pd specimens which were charged by cathodic hydrogen, with the micro Vickers hardness tester, and have obtained several interesting results in regard to hydrogen hardening processes.

Figure 1 shows the relation between the hardness of Pd and the duration of hydrogen charging, where the curve *a* is for the perfectly annealed (1200°C−3 hrs) specimen, *b*, *c*, and *d* for the lightly, mediumly, and severely cold worked specimen, respectively. In the case of cathodic charging by the current density of 3.6 mA/cm², it is shown from the electric resistance measurement that the α phase exists up to 10 min of charging and thereafter it transforms to the β phase, and also shows that some curves have maxima and others do not. All curves in Fig. 1 show the increasing hardness during the initial ten min of charging. When hydrogen was charged long enough, the hardness reached a definite value. We know, therefore, that the increasing part of the hardening curve corresponds to the α phase and the part of a definite value to the β phase. The final hardness value of the severely cold worked specimen is nearly equal to the initial value, but other specimens have higher values of the final hardness than the initial.

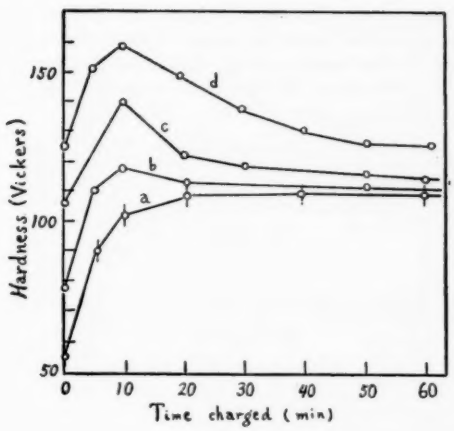


FIG. 1.

Figure 2 shows the change of hardness of the hydrogen-charged Pd specimen with discharging times, where the reversed electrolysis is used for discharging. The initial state of specimens is the perfectly annealed. In the case of the α phase, the hardness decreases monotonously and quickly reaches its initial value, while the curve of the β phase once ascends with discharging times and reaches a definite value (higher than the initial value) through a maximum after a long time.

From these experimental results, it becomes evident that the α and the β phase may have different mechanisms of hydrogen hardening. The hardness of the α phase increases when the hydrogen content increases, and reversibly decreases with decreasing hydrogen content. After discharging had been finished, the specimen was again heated at 1200°C and microscopically investigated. We could not detect any observable change in its structure. We know, therefore, that the hydrogen hardened α state completely recovers to its initial annealed state only by discharging the hydrogen.

The hardening mechanism of the β phase may be somewhat complicated. When the specimen was charged with hydrogen to a β range, the slip-band-like traces appeared in the grain structures. These traces remained after the hydrogen had been perfectly discharged. The reason why the hydrogen hardened β phase specimen does not recover its initial softness after discharging hydrogen can be explained by these traces. After complete discharging, in this case, the hydrogen hardened Pd could be

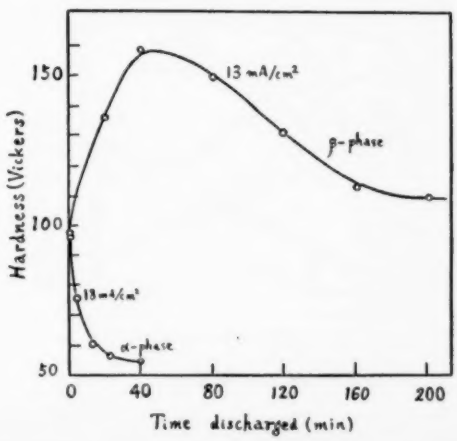


FIG. 2.

recrystallized at 700°C. The essential nature of these traces and of the β phase are now under investigation. Some tentative explanation about the mechanism of hydrogen hardening phenomena in a β range and the details of experimental results will be shown in another paper.[3]

[1] D. P. Smith, *Hydrogen in Metals* (University of Chicago, Chicago, 1948).
[2] T. Sugeno and M. Kowaka, Mem. Inst. Sci. and Ind. Research, Osaka Univ. **XI** (1954).

Elastic Relaxation and Grain Boundary Diffusion in Cadmium

Y. L. Yousef and R. Kamel
Physics Department, University of Cairo, Giza, Egypt
(Received April 7, 1954)

INTERNAL friction and Young's modulus relaxation, in relation to intergrain diffusion, were investigated for pure cadmium in the polycrystalline and in the single crystal states. The specimens were in the form of short wires contained in an evacuated tube which could be maintained at various temperatures. Transverse vibrations with amplitudes of the order of 10^{-4} cm at low sonic frequencies were set up electrostatically and were detected by a capacity-responsive circuit which has earlier been described in connection with sodium ion diffusion in glass.[1] The internal friction Q^{-1} was measured by the resonance curve method. Young's modulus E was determined at various temperatures simultaneously with the measurement of internal friction. The square of the frequency of vibration was taken as a measure of Young's modulus.

In conformity with the poineer work of Kê[2] on aluminum by a torsion pendulum method, the internal friction *versus* temperature curve for polycrystalline cadmium showed a marked absorption band due to relaxation by intergrain diffusion. A relaxation in Young's modulus *versus* temperature curve manifested itself as an inflection in the curve, and occurred in the same temperature range where the absorption band appeared (Fig. 1). In single crystal specimens prepared by the moving furnace method,[3] a regular behavior was observed, indicating the absence of any relaxation process under the experimental conditions contemplated here. Since the observed absorption band and the relaxation of the elastic modulus in the polycrystalline cadmium disappeared completely in the case of single crystal specimens, the mechanism responsible for the present anelastic effect in polycrystals must be the diffusion of grain boundaries.

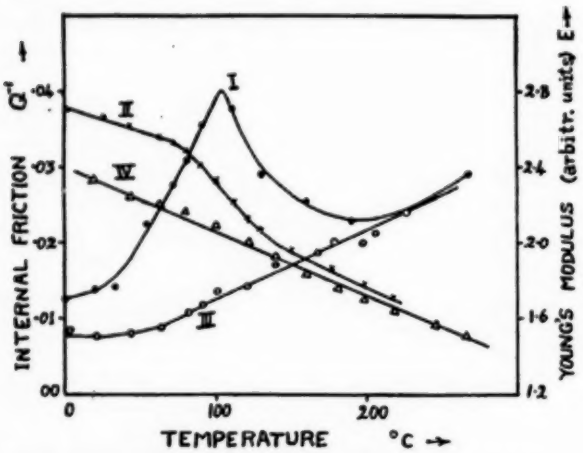


FIG. 1. Internal friction (●,○) and elasticity relaxation (×,△) curves for cadmium in the polycrystalline (I and II) and in the monocrystalline (III and IV) states. The frequency at room temperature for curves I and II is 41.5 c/s and for curves III and IV is 48 c/s.

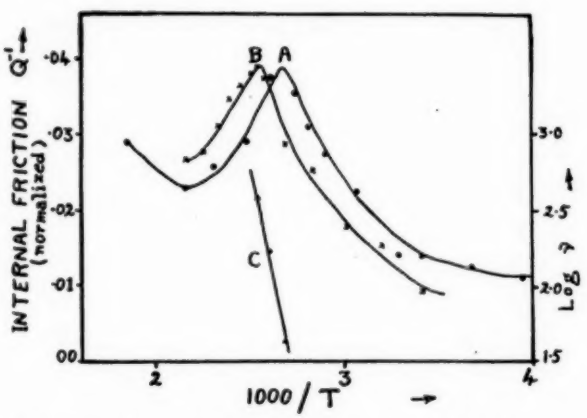


FIG. 2. Effect of change of temperature on the position of the internal friction peak of two cadmium strips (8.35 × 0.125 × 0.078 cm and 2.65 × 0.125 × 0.078 cm). The frequency at room temperature for curve A is 41.5 c/s and for B is 385 c/s. Curve C represents the position of the peak as a function of frequency.

The heat of activation H associated with grain boundary diffusion was deduced from Fig. 2 and was found to be 32 kcal/mole. According to Le Claire,[4] the value of H may be taken, without appreciable error, as equal to its value for self-diffusion. So far as we are aware, it seems that no results were previously published on the self-diffusion of cadmium.

[1] R. Kamel, J. Appl. Phys. **24**, 1308 (1953).
[2] T. S. Kê, Phys. Rev. **71**, 533 (1947); **72**, 41 (1947); **73**, 267 (1948).
[3] E. N. da C. Andrade and R. Roscoe, Proc. Phys. Soc. (London) **49**, 152 (1937).
[4] A. D. Le Claire, Phil. Mag. **42**, 468 (1951).

The Measurement of Electromagnetic Field in Resonant Cavities by Introduction of Small Metallic Rings

Aldo Gilardini
Istituto Superiore PP.TT., Rome, Italy
(Received April 5, 1954)

IN a recent paper[1] published in this journal, L. C. Maier and J. C. Slater have shown that the electromagnetic field in a resonant cavity can be determined by measuring the perturbation of the resonant frequency, caused by the insertion of a small metallic element at the point where the electromagnetic field has to be measured. The frequency shift is given by an equation which can be stated in the form[2]

$$\frac{f_0^2-f^2}{f_0^2}=\frac{1}{2\mu}[\epsilon(M_xE_x^2+M_yE_y^2+M_zE_z^2) -\mu(N_xH_x^2+N_yH_y^2+N_zH_z^2)],$$

where f_0 is the unperturbed resonant frequency, f the perturbed frequency, U the total electromagnetic energy stored in the cavity, E_x, E_y, E_z, H_x, H_y, H_z, the six components of the electromagnetic field at the point of perturbation in an arbitrary Cartesian coordinate system, M_x, M_y, M_z, N_x, N_y, N_z parameters having volume dimensions, related to the size and shape of the element and to its orientation in the coordinate system. (It is interesting to note that ϵM_iE_i and $-\mu N_iH_i$ are the components along the i axis of the electric and magnetic moment of the perturbing element.)

In order to determine the six components of the electromagnetic field a system of six equations is required; this system can be obtained by six frequency shift measurements, each with different shape or orientation of the perturbing element, supposing M and N parameters known.

Maier and Slater have calculated these parameters for an ellipsoid of revolution and found good agreement with experiment

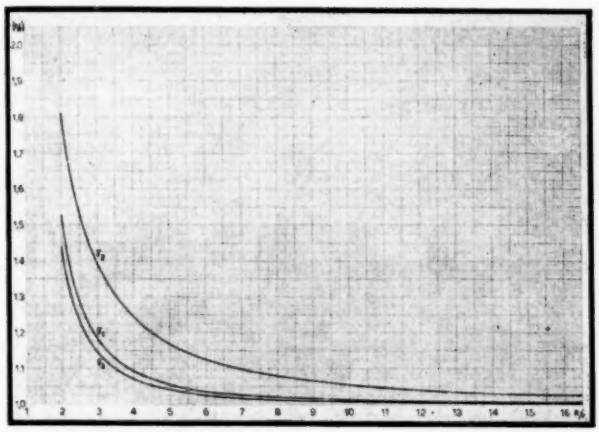


FIG. 1.

in the extreme cases of spheres and very thin disks. In the general case the agreement is much less satisfactory, probably, as the authors suggest, owing to the poor tolerances in machining an ellipsoid.

By using spheres and very thin disks with different orientations obviously only four independent equations can be obtained and for the remaining two, general ellipsoids must be used. In order to improve the precision I have tried to substitute to the ellipsoid, in the last two measurements, another metallic element of a shape which can be more easily machined. A toroidal ring seems easier to machine and I have considered the possibility of using it.

Let us assume that the rotation axis of a toroidal ring is directed along the z coordinate axis, and let r be the radius of the generating circle and R the radius of the circumference described by the center of the generating circle. M and N parameters can be calculated following the method illustrated by Maier and Slater for ellipsoids. The results are[3]

$$M_x=4\pi^2\frac{(R^2-r^2)^{\frac{1}{2}}}{(R/r)^2}F_1\left(\frac{R}{r}\right)$$

$$M_x=M_y=2N_z=2\pi^2\frac{(R^2-r^2)^{\frac{1}{2}}}{\ln(8R/r)-2}F_2\left(\frac{R}{r}\right)$$

$$N_z=N_y=3\pi^2\frac{(R^2-r^2)^{\frac{1}{2}}}{(R/r)^2}F_3\left(\frac{R}{r}\right)$$

where F_1, F_2, and F_3 are functions of R/r, which have unity asymptotic values when $R/r\to\infty$. F functions can be given as expressions containing associated Legendre functions of semi-integer degree and have been calculated and plotted in Fig. 1 for $R/r\geqslant 2$.

It is interesting to note that the relationship $M_x=M_y=2N_z$ is also valid for ellipsoids.

In Table I are reported the results of an experimental verification. Two tori have been used, the first with $R=3.65$ mm, $r=0.73$ mm, the second with $R=3.09$ mm, $r=1.51$ mm. They have been inserted in a rectangular cavity, operating in the TE_{106} mode ($f_0=3040$ MHz), at the points of maximum electric field and of maximum magnetic field. Estimated error in Δf is $\pm 10kHz$, which is larger than Maier and Slater's inaccuracy, owing to a different method of measuring Δf, but is good enough for our purposes. The percentage errors reported in the last column show that perturbation theory for toroidal rings checks to within a few percent, as Maier and Slater already found for spheres and very thin disks.

TABLE I.

Torus	Parameter	Δf (MH_z)	Experimental value of parameter (mm^3)	Calculated value (mm^3)	Percentage of error
1°	M_x	−3.91	642	627	2.4
1°	M_z	−0.48	79	76.5	3.3
1°	N_x	+0.19	59	56.6	4.2
1°	N_z	+0.98	303	313	3.2
2°	M_x	−5.41	888	867	2.4
2°	M_z	−1.66	272	275	1.1
2°	N_x	+0.62	192	195.5	1.8
2°	N_z	+1.36	420	434	3.2

The writer wishes to express his hearty thanks to Professor V. Gori, Director of the Istituto Superiore PP.TT., for his kind encouragement and support in this study.

[1] L. C. Maier, Jr., and J. C. Slater, J. Appl. Phys. **23**, 68 (1952).
[2] H. B. G. Casimir, Philips Research Reports **6**, 162 (1951).
[3] Calculations are developed in a report of the Istituto Superiore PP.TT., which the author will be glad to send on request. Requests can be addressed to him: Viale Trastevere 189, Rome, Italy.

Centripetal and Centrifugal Electromagnetophoresis

ALEXANDER KOLIN*
University of Chicago, Chicago, Illinois
(Received April 7, 1954)

IT has been reported in a previous paper[1] that particles suspended in a liquid whose electrical conductivity differs from that of the particles experience a force when a current is passed through the liquid at right angles to a magnetic field which permeates the liquid. To avoid convection of the fluid, the current, as well as the magnetic field, must be homogeneous. The magnetic field of flux density B and the current i may either be constant or alternating; in the latter case a maximum force is obtained when B and i are in phase. The direction of the force $\mathbf{F}$ is perpendicular to $\mathbf{B}$ and $\mathbf{j}$ ($\mathbf{j}$ being the current density in the suspension). Particles whose conductivity (σ') exceeds that of the liquid (σ'') experience a force in the direction of the Lorentz forces exerted upon the moving ions, whereas particles of $\sigma'<\sigma''$ experience a force in the opposite direction. Particles of $\sigma'=\sigma''$ experience no force.

The phenomenon is the result of a combination of body forces and surface forces exerted upon suspended particles subjected to the above-mentioned conditions. The body force is an electromagnetic force exerted upon the current passing through the particle; the surface force is a combination of hydrostatic forces (brought about by a hydrostatic pressure field set up in the fluid in the electromagnetic field) with hydrodynamic forces manifested as a result of fluid streaming engendered in the fluid adjacent to the particle. The situation is roughly analogous to particles suspended in a field of gravity, where the electromagnetic body force is the analog of the weight and the surface force is that of the buoyancy.

A theoretical estimate of the force upon small spheres, taking viscous drag into account,[2] predicts a force

$$\mathbf{F}=\frac{3}{2}[\mathbf{B}\times\mathbf{j}]V\left[\frac{\sigma''-\sigma'}{2\sigma''+\sigma'}\right], \tag{1}$$

where V is the volume of the sphere.

For nonconducting spheres suspended in a conducting fluid ($\sigma'\ll\sigma''$) one obtains, thus,

$$\mathbf{F}=\frac{3}{4}[\mathbf{B}\times\mathbf{j}]V. \tag{2}$$

This expression has been verified by observing the rate of migration of polystyrene spheres of diameters in the order of magnitude of 100 μ suspended in NaCl solutions.

The main experimental difficulty in observing this effect consists in the exclusion of electromagnetically engendered convection. The slightest nonuniformity in the magnetic field or current density may give rise to liquid convection which could mask the effect. The cell containing the suspension must, therefore, be made with great precision so as to minimize variations in the

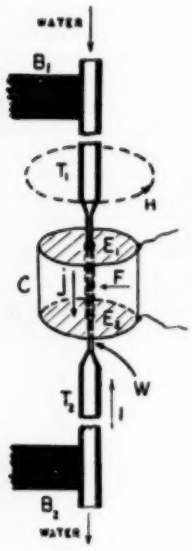


FIG. 1. Generator of solenoidal magnetic field and cylindrical electromagnetophoresis cell C. W: Thin water-cooled copper tubing. T_1 and T_2: Heavier copper tubing ($\frac{1}{4}$ in. od; $\frac{1}{8}$ in. id) conveying the current to W. B_1 and B_2: Heavy copper bars (1 in. in diameter, 32 in. long connecting T_1 and T_2 to the secondary of the stepdown transformer. (Distance between B_1 and B_2: 30 in. Length of W: 10 in.). E_1 and E_2: Electrodes. j: Current density vector parallel to W. F: Vector representing the centripetal force exerted upon suspended nonconducting particles.

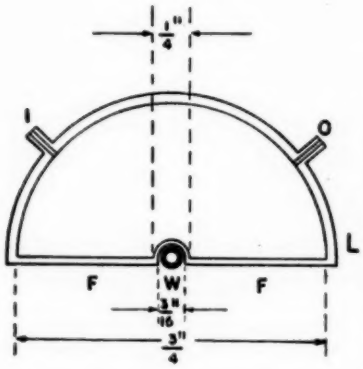


FIG. 2. Cross section of semicylindrical electromagnetophoresis cell. L: Lucite "washer," a transparent wall (5 mm high) sandwiched between two silver electrodes. (The electrodes are parallel to the plane of the drawing and are not shown in the figure. They are cemented, for greater rigidity, to Lucite plates.) I and O: Channels through which the cell can be filled. F: Flat walls through which the migration of the particles toward W or away from it may be observed. W: Copper tubing which serves as electromagnet.

cross section perpendicular to **j** and the greatest care must be taken to maximize the uniformity of the magnetic field.

Because of the practical difficulties involved in obtaining intense magnetic fields of great uniformity, it appeared desirable to investigate the possibility of observing the effect with nonuniform fields. This is indeed possible by using combinations of electric and magnetic fields which satisfy the condition curl $[\mathbf{j}\times\mathbf{B}]=0$. A field combination which appears to be especially favorable consists of the solenoidal magnetic field surrounding a very long wire (carrying a current I) with a current of uniform density **j** traversing the cell parallel to the wire (Fig. 1). The electromagnetic forces exerted upon the fluid as well as upon the particles, being perpendicular to both **B** and **j**, are radial. Thus, suspended particles experience centripetal or centrifugal forces, depending on the phase relation between **B** and **j**. For particles which are small as compared to the radius of the wire W, Eq. (1) can be assumed to hold with a good degree of approximation.

The arrangement shown in Fig. 1 enables one to obtain intense magnetic fields within a small volume adjacent to the surface of W, which is adequate for work with small particles. Currents of several thousand amperes can be readily obtained with powerful stepdown transformers. Magnetic fields of sufficient symmetry can be obtained by making wire W part of a very large rectangular loop which is energized by the stepdown transformer. The "magnet" W can be conveniently cooled by using copper tubing, rather than a solid wire. In our experiments currents of up to about 3500 amperes were passed through a copper tubing of $\frac{3}{16}$ in. outside diameter and $\frac{3}{32}$ in. inside diameter. Thus, the value of **H** at the surface of the wire was approximately 2800 oersteds. The tubing W was cooled by tap water passing through it at the approximate rate of 12 liters/minute. **j** varied from 0.1 amp/cm² to 1 amp/cm². For example, suspensions of polystyrene spheres of average diameter of 115μ (ranging from 25μ to 400μ) have been "centrifuged" successfully. At a current of $I=3050$ amp passing through the tubing W with a current of density $\mathbf{j}=0.65$ amp/cm² traversing the suspension, the precipitation around the central tubing W was nearly complete (except for a few spheres of the smallest size) within 5 seconds and practically complete within 10 seconds. In dense suspensions the deposit of particles will be heavy and not uniform. The nonuniform deposit creates an electromagnetic convection in its vicinity which tends to redistribute the particles uniformly so that the convection slows down and finally stops.

It is not necessary to use a cylindrical cell surrounding W as shown in Fig. 1. Cells of a great variety of forms can be used. A semicylinder, or better, a cell of the cross section shown in Fig. 2 is particularly convenient. W indicates the location of the wire which generates the magnetic field. The transparent Lucite "washer" L shown in Fig. 2 is sandwiched between two similarly shaped silver electrodes. The migration of the particles toward the center of the wire W or away from it can be easily observed through the flat Lucite walls F of the cell. It is important to align the planes of the electrode, E_1 and E_2 (Fig. 1) exactly perpendicular to the wire W in order to minimize electromagnetic convection. Thin silver electrodes (3/1000 in.) have been used to avoid excessive heating of the electrodes by eddy currents thus reducing thermal convection.

The author wishes to express his indebtedness to the Research Corporation and to the Smith, Kline and French Foundation for generous support of his research. Thanks are also due Mr. Edward Wolowiec for technical assistance.

* This work has been supported by a grant from the U. S. Office of Naval Research.

[1] A. Kolin, Science **117**, 137 (1953).
[2] D. Leenov and A. Kolin, J. Chem. Phys. **22**, 683 (1954).

On Shock Propagation in Brass

H. BECKER*
Nuclear Power Division, Combustion Engineering, Inc., New York, New York
(Received May 20, 1953)

IN a recent paper, Shreffler and Deal[1] (SD) presented velocity-thickness data for the free surface motion of brass plates under explosive detonation. Savitt[2] (S) plotted the reciprocal of the initial free surface velocity from (SD) as a function of plate thickness, and obtained a sharp discontinuity at about $\frac{1}{4}$ in.

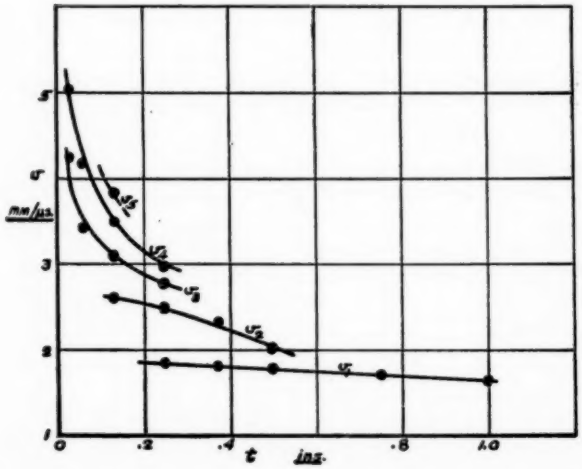


FIG. 1. v–t plot.

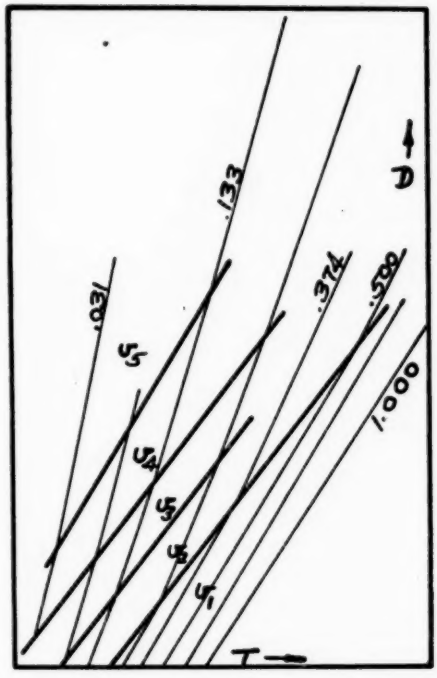


FIG. 2. Velocity zones.

thickness. Actually, a clearer picture of the behavior of the plates may be obtained by plotting all (SD) velocities against plate thickness. This has been done in Fig. 1, from which it is apparent that several curves may be drawn connecting the velocity values, thus yielding lines of demarcation for velocity zones in the plates. This has been done in Fig. 2 for which the fine lines were taken from Fig. 5 (SD). The demarcation zones are clearly indicated, thus revealing cut-off thickness values at which velocity ranges in the thicker plates are no longer possible in the thinner plates. It would appear, from the sharp slope changes of the (SD) plots, that the velocity zones are defined by specific deflection modes, some of which may exist in thick plates but may not exist in thin plates. Furthermore, it could be presumed that a sharp change in slope of the line for any particular thickness may be due to a change in mode shape.

* Now at New York University, Engineering Research Division, University Heights, New York. Revised manuscript received on May 20.
[1] R. G. Shreffler and W. G. Deal, J. Appl. Phys. **24**, 44 (1953).
[2] J. Savitt, J. Appl. Phys. **24**, 1335 (1953).

On the Conduction of Heat into a Growing Vapor Bubble

H. KURT FORSTER
Department of Engineering, University of California, Los Angeles, California
(Received April 16, 1954)

IN a recent paper on the integration of Rayleigh's equation for the problem of a vapor bubble growing in a superheated liquid we used a simple expression for the temperature of the moving bubble wall with very good success.[1] An exact analysis is forthcoming but in view of possible further applications to related problems, particularly to the problem of the collapsing bubble (cavitation) and the bubble growing under the influence of electric charges due to ionizing radiation the detailed considerations which led us to our formulation may be of interest.

The two-phase heat conduction problem with phase boundary moving in consequence of the phase change is quite old and was first solved for some special cases by Franz Neumann as reported in Riemann's lectures.[2] Carslaw and Jaeger who also substantially contributed to this problem furnish a rather complete list of references.[3] A paper dealing with the case of a moving spherical boundary between liquid and vapor was published a few years ago by Plesset and Zwick,[4] who assumed a "thermal boundary layer" around the vapor bubble and obtained a first approximation for the bubble temperature in a method of successive approximations. In our present treatment, we use Lightfoot's[5] idea of representing the heat liberated or absorbed in phase change by sources or sinks, but instead of solving the heat conduction and (heat) convection problem at once and together we solve the more important of the two problems (heat conduction) first and exactly, and then, having a general solution of the head conduction problem valid for all bubble radii, we investigate the influence of the radial velocity on the temperature at the phase interface.

Our system consists of a vapor bubble of radius R_0 embedded in an infinite fluid. Specific heat, thermal conductivity and diffusivity and heat of evaporation will be denoted by c, K, κ, and L and subscripts one and two will indicate the vapor or liquid phase, respectively. Initially, liquid and vapor are at the same temperature (boiling point plus superheat) which is taken as the zero of our temperature scale.

If the radius of the bubble increases by dR_0, an amount of heat $dQ=L\rho_1 4\pi R_0^2 dR_0$ is absorbed constituting a negative distributed spherical surface heat source. The heat of evaporation is so large that its influence is the only important contribution toward the subsequent temperature distribution. To orient ourselves about the nature of the problem we shall first disregard the bubble wall curvature; this is a useful consideration because the essential features can be seen very clearly for the plane interface and, as will be shown later, consideration of curvature merely introduces some rather simple changes.

We assume the phase interface in the (y,z) plane with vapor for $x<0$ and liquid for $x>0$; the initial condition is $T_1=T_2=0$ except for $x=x'$ and the boundary conditions are

$$T_1=T_2 \text{ and } K_1\frac{\partial T_1}{\partial n_1}=K_2\frac{\partial T_2}{\partial n_2} \text{ at the interface.} \tag{1}$$

The temperature distribution due to an instantaneous distributed plane source $dQ'=dQ/4\pi R_0^2$ at $x=x'$ is then[6]

$$dT_2(x,t)=\frac{dQ'}{2(\pi\kappa_2 t)^{\frac{1}{2}}}\Big[\exp(-\beta(x-x')^2)+\frac{K_2\kappa_1^{\frac{1}{2}}-K_1\kappa_2^{\frac{1}{2}}}{K_2\kappa_1^{\frac{1}{2}}+K_1\kappa_2^{\frac{1}{2}}}\exp(-\beta(x+x')^2)\Big], \tag{2}$$

where $\beta=(4\pi t)^{-1}$ [we should correctly divide by $c_2\rho_2$; lest the formulas become too cumbersome we have done so only in our final Eqs. (9) and (10)]. In our case the heat source dQ' is at the interface, thus $x'=0$; and in order to obtain the temperature at the bubble wall (which we shall always denote by T without subscript) we set $x=0$, giving

$$dT=\frac{dQ'}{(\pi\kappa_2 t)^{\frac{1}{2}}}\cdot\frac{1}{1+\frac{K_1}{K_2}\left(\frac{\kappa_2}{\kappa_1}\right)^{\frac{1}{2}}}. \tag{3}$$

The constant factor which appears in Eq. (3) is very close to 1, e.g. for water at about 100°C it is $(1.0032)^{-1}$. Its physical significance may be found in the fact that the ratio $K_1\kappa_2^{\frac{1}{2}}/K_2\kappa_1^{\frac{1}{2}}:1$ (which is here numerically about 3:1000) gives the proportion of heat flowing into the source at the bubble wall from the vapor region to that coming from the region filled with liquid.

If we now consider the (small) spherical bubble, this ratio will become even smaller because the total heat contained in the extremely small initial bubble (its radius is about 10^{-3} cm for a few degrees of superheat) is vanishingly small and then, as the bubble grows, the vapor emanating from the bubble wall by evaporation is below the initial temperature of the system. We shall therefore choose as our boundary condition at the bubble wall of the spherical bubble: no heat flux across the phase interface.

Our problem is therefore to find the temperature distribution in the liquid due to a spherical surface heat source of strength 1 at R' with the boundary condition

$$\left(\frac{dT_2}{dR}\right)_{R=R_0}=0; \quad R_0 \leq R'; \quad T_2 \text{ finite at infinity.} \tag{4}$$

The solution of the problem is the Green's function for the domain for the stated conditions; it does not seem to be given in the literature and we derived it along lines given in various books for similar problems.[7] The Laplace transform $\bar{T}_2$ of the temperature in the liquid T_2 is obtained as

$$\bar{T}_2=(8\pi RR'\kappa_2 q)^{-1}\left[\exp(-q|R-R'|)+\frac{qR_0-1}{qR_0+1}\exp(q(2R_0-R'-R))\right], \tag{5}$$

for $R_0 \leq R \leq R'$; q stands for $(p/\kappa_2)^{\frac{1}{2}}$, and p is the transformation variable in the usual notation. The inverse transform of Eq. (5), which may be obtained from tables, yields the temperature in the liquid. In our case we wish to know the temperature at the phase interface (and therefore we let R approach R_0) when evaporation takes place at the interface ($R' \to R_0$) which gives us the transform of the bubble wall temperature due to a unit heat source

$$\bar{T}=\lim_{R'\to R_0}\lim_{R\to R_0}\bar{T}_2=[4\pi R_0\kappa_2(qR_0+1)]^{-1}. \tag{6}$$

The temperature of the bubble wall of a bubble of radius R_0 due to a source dQ at R_0 then becomes

$$dT=\frac{dQ}{4\pi R_0^2(\pi\kappa_2 t)^{\frac{1}{2}}}-\frac{dQ\operatorname{erfc}[(\kappa_2 t)^{\frac{1}{2}}/R_0]\exp(\kappa_2 t/R_0^2)}{4\pi R_0^3}. \tag{7}$$

We note that in consequence of the factor $t^{\frac{1}{2}}$ in the denominator the important contributions to the temperature take place in the neighborhood of $t=0$; in fact, if we take the ratio of the second to the first term in Eq. (7) and write x for the dimensionless quantity $(\kappa_2 t)^{\frac{1}{2}}/R_0$ we see that it is

$$\varphi(x)=\pi^{\frac{1}{2}}x\,e^{x^2}\operatorname{erfc}x, \tag{8}$$

which is zero for $x=0$. Let us further note that the first term in Eq. (7) is identical with the solution for the plane given by Eq. (3) for $K_1=0$, and that the ratio given by Eq. (8) remains small for longer times as we take larger radii.

If we now consider the influence of the phase boundary motion on the bubble wall temperature we see directly that for the plane this influence is nil, i.e., this temperature is given by Eq. (3) or the first term of Eq. (7) whether the phase boundary moves or not.

The physical significance of the dimensionless quantity x may be seen from the fact that $(\kappa_2 t)^{\frac{1}{2}}$ is essentially the distance traversed by the temperature wave* and x is therefore this distance over the bubble radius. This consideration gives us a means to distinguish between slow and fast growing bubbles, depending on whether the bubble growth velocity is smaller or larger than the velocity of the temperature wave. In our case, for a few degrees of superheat[1] the bubble growth velocity is larger than the velocity of the temperature wave by a factor of about ten. More detailed considerations[8] show that x, starting from zero will remain smaller than about 0.05 during our time interval of interest which is less than 0.1 second. The ratio $\varphi(x)$ then remains smaller than 0.1 and we may neglect the second term in Eq. (7) compared to the first since we thereby introduce a small error in a small quantity.

Concerning the influence of radial motion of the bubble wall on the first term in Eq. (7) we first note that the bubble, if left at radius R_0 will at time t have a temperature $\sim 1/R_0^2$; if, however, immediately after liberation of dQ the bubble is expanded to radius R, then its temperature at time t will be $\sim 1/R^2$ [and it will be given by the first term of Eq. (7) with R_0 replaced by R]. In reality the bubble wall has in time t moved from R_0 to R and a mean value of $1/R^2$ inserted in Eq. (7) should furnish a good approximation to the actual temperature. To obtain such a mean value we integrate $1/R^2$ from R_0 to R and divide by the interval; this gives $1/RR_0$. This value, inserted into the first term of Eq. (7) together with the expression for dQ as given previously yields

$$dT=\frac{L\rho_1 R_0 dR_0}{c_2\rho_2 R(t)(\pi\kappa_2 t)^{\frac{1}{2}}}. \tag{9}$$

The temperature which the bubble attains in consequence of the evaporation taking place continuously while it grows from R_0 to R is now simply the sum over all the contributions given by Eq. (9)

$$T=\frac{L\rho_1}{c_2\rho_2(\pi\kappa_2)^{\frac{1}{2}}}\int_0^t\frac{R(x)\dot{R}(x)dx}{R(t)(t-x)^{\frac{1}{2}}}. \tag{10}$$

Equation (10) gives the temperature of a (fast) growing spherical vapor bubble whose center is at rest in an infinite liquid.

The temperature integral just derived is identical with the expression used in the treatment of the bubble growth problem in reference 1. It is interesting to point out that the bubble growth problem is quite insensitive to the exponent of (R_0/R) in Eq. (10). In fact, if $[R(x)/R(t)]^n$ (n a positive integer) is used in Eq. (10), the exponent n (a) leaves the solution of the bubble growth integral equation in the vicinity of $t=0$[9] absolutely unaffected and (b) leaves this solution for large times unaffected in its form (Constant $\cdot t^{\frac{1}{2}}$), merely changing the constant by a factor, which is, moreover, of order unity for any reasonable n. This weak dependence on the exponent n is a consequence of the occurrence of $(t-x)^{\frac{1}{2}}$ in the denominator of Eq. (10) whereby the contributions in the neighborhood of $x \approx t$ [i.d., when $R(x)/R(t) \approx 1$] are much more important than the subsequent contributions.

[1] H. K. Forster and N. Zuber, J. Appl. Phys. **25**, 474 (1954).
[2] Riemann-Weber, *Die Partiellen Differential Gleichungen der Mathematischen Physik* (Vieweg und Sohn, Braunschweig), fifth edition, pp. 117–121.
[3] H, S, Carslaw and J. C. Jaeger, *Conduction of Heat in Solids* (Oxford University Press, London) first edition, p. 71. This work served as a general reference in the present study.
[4] M. S. Plesset and S. A. Zwick, J. Appl. Phys. **23**, 95 (1952).
[5] Lightfoot, Proc. London Math. Soc. (2) **31**, 97 (1929).
[6] See reference 3, p. 300.
[7] See for instance, reference 3, chapter XII.
* Let the temperature have a maximum at R_0 at $t=0$; then a temperature maximum will occur at $R_0+(2\kappa t)^{\frac{1}{2}}$ after time t; we shall call the velocity of this disturbance the velocity of the temperature wave.
[8] See reference 1, Eqs. 13 and 14.
[9] See reference 1, Eq. 11.